

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:23:35 ; Search time 412.25 Seconds
(without alignments)
11544.662 Million cell updates/sec

Title: US-09-898-556a-3
Perfect score: 2772
Sequence: 1 caggcgcttaagctggtg.....ttctaccatctctaccct 2772

Scoring table: OLIGO_MNC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database : N.Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	56.9	2614	23	AAST4823 DNA encoding novel
2	571	20.6	678	22	AAAF4909 Human breast cancer
3	525	18.9	653	23	AAAS68256 DNA encoding novel
4	231	8.3	983	22	AAAS26286 Human cDNA encoding
5	99	3.6	394	22	AAI83890 Human polynucleoti
6	69	2.5	69	22	AAI13031 DNA encoding zinc
7	69	2.5	69	22	AAAS13088 DNA encoding novel
8	49	1.8	821	23	AAAS68257 Human colon cancer
9	29	1.0	800	22	AAH33832

10	29	1.0	2349	23	AAAS90429 DNA encoding novel
11	29	1.0	3582	22	AAIS8664 Human polynucleoti
12	29	1.0	3582	22	AAI60450 Human polynucleoti
13	28	1.0	159	22	AAAL00216 Human reproductive
14	27	1.0	1183	22	AAAS25949 Human cDNA encoding
15	27	1.0	2717	22	AAAS26635 Human genomic DNA
16	27	1.0	3099	20	AAAV64361 Human stem cell z1
17	25	0.9	51	23	ABLU00474 Human silent nonco
18	25	0.9	51	23	ABLU00878 Human amino acid c
19	25	0.9	763	22	AAK92290 Human cDNA 5'-end
20	25	0.9	763	22	AAK93811 Human cDNA clone r
21	25	0.9	814	22	AAI95339 Human neuroblastom
22	25	0.9	871	22	ABAO8823 Human PRO1847 homo
23	25	0.9	2200	22	AAK94575 Human full-length
24	23	0.8	688	22	AAK91560 Human cDNA 5'-end
25	23	0.8	688	22	AAK93129 Human cDNA clone r
26	23	0.8	1512	22	ABA50438 Human breast cell
27	23	0.8	1512	22	ABA68388 Human foetal liver
28	23	0.8	1512	22	ABA35382 Probe #13848 for g
29	23	0.8	1512	22	AAK16761 Human brain expres
30	23	0.8	1512	22	AAK42537 Human bone marrow
31	23	0.8	1512	22	AAI23284 Probe #13217 for g
32	23	0.8	1512	22	AAI48604 Probe #17290 used
33	23	0.8	1512	22	AAI08925 Probe #8916 used t
34	23	0.8	1623	22	AAK94635 Human full-length
35	23	0.8	1965	22	ABA45305 Human breast cell
36	23	0.8	1965	22	ABA55794 Human foetal liver
37	23	0.8	1965	22	ABA25474 Probe #3940 for ge
38	23	0.8	1965	22	AAK04015 Human brain expres
39	23	0.8	1965	22	AAK29501 Human bone marrow
40	23	0.8	1965	22	AAI14072 Probe #4005 for ge
41	23	0.8	1965	22	AAI35452 Probe #4138 used t
42	23	0.8	1965	22	AAI03925 Probe #3916 used t
43	23	0.8	2534	23	AAAS81041 DNA encoding novel
44	22	0.8	439	22	AAAF67397 Novel human polynu
45	22	0.8	484	21	AACT92334 Human lung tumour-
46	22	0.8	484	21	AAAD23110 Human lung tumour-
47	22	0.8	758	22	AAAS26521 Human cDNA encoding
48	22	0.8	1038	22	ABA21456 Human nervous syst
49	22	0.8	1038	22	AAK84268 Human immune/haema
50	22	0.8	1185	23	AAAS71465 DNA encoding novel
51	22	0.8	1194	23	AAAS67440 DNA encoding novel
52	22	0.8	1664	23	AAAS64739 Human polynucleoti
53	22	0.8	1461	22	AAK51912 Human polynucleoti
54	22	0.8	1599	23	AAAS71467 DNA encoding novel
55	22	0.8	1812	22	AAK52682 Human polynucleoti
56	22	0.8	2072	22	AAK52896 Human polynucleoti
57	22	0.8	2199	22	AAK51698 Human polynucleoti
58	22	0.8	2259	22	AAAS26098 Human cDNA encoding
59	22	0.8	2382	20	AAAV64360 Human stem cell z1
60	22	0.8	2391	22	AAK94745 Human full-length
61	22	0.8	2472	23	AAAS71466 DNA encoding novel
62	22	0.8	2659	22	AAK52096 Human polynucleoti
63	22	0.8	2659	22	AAK52096 Human transcriptio
64	22	0.8	4029	23	AAAS65913 Human genomic DNA
65	22	0.8	11221	22	AAAS26799 Human melanog
66	21	0.8	264	23	ABLI21853 Drosophila melanog
67	21	0.8	300	21	AAAO0606 Human colon cancer
68	21	0.8	455	23	AAAS86928 DNA encoding novel
69	21	0.8	477	22	AAI92157 Human polynucleoti
70	21	0.8	584	22	AAAL14925 Human breast cancer
71	21	0.8	605	22	AAH07232 Human cDNA clone (
72	21	0.8	693	22	AAH04531 Human cDNA clone (
73	21	0.8	752	21	AAA02263 Human colon cancer
74	21	0.8	781	22	AAI23783 Human breast cancer
75	21	0.8	848	22	AAAS31025 Human diagnostic a
76	21	0.8	874	22	AAH06905 Human cDNA clone (
77	21	0.8	1426	22	AAK52270 Human polynucleoti
78	21	0.8	1494	24	ABR98977 Human cancer supp
79	21	0.8	1663	20	AAK39662 Renal cancer assoc
80	21	0.8	1705	22	AAK53253 Human polynucleoti
81	21	0.8	1705	22	AAK53254 Human polynucleoti
82	21	0.8	1765	24	AAAS18786 cDNA encoding huma

83	21	0.8	1828	22	AAK52269	Human polynucleoti
84	21	0.8	1984	22	AAK94879	Human full-length
85	21	0.8	2008	22	ABA09199	Human NY-REN-21 Ag
86	21	0.8	2008	22	AAI60805	Human polynucleoti
87	21	0.8	2038	22	AAI59019	Human polynucleoti
88	21	0.8	2377	23	ABU21852	Drosophila melanog
89	21	0.8	2382	22	AAI66707	A. gossypii Ag011
90	21	0.8	2732	22	AAH17926	Human cDNA sequenc
91	21	0.8	2735	22	ABU21824	Drosophila melanog
92	21	0.8	3204	22	AAH17368	Human cDNA sequenc
93	21	0.8	4199	23	ABU22380	Drosophila melanog
94	21	0.8	4829	23	AAK86168	DNA encoding novel
95	21	0.8	6930	23	AAK86280	Human immune/haema
96	21	0.8	10926	22	AAK65370	Human immune/haema
97	20	0.7	127	22	ABA49753	Human breast cell
98	20	0.7	127	22	ABA67665	Human foetal liver
99	20	0.7	127	22	ABA75338	Human foetal liver
100	20	0.7	127	22	ABA34733	Probe #13199 for g
101	20	0.7	127	22	AAK16083	Human brain expres
102	20	0.7	127	22	AAK41825	Human bone marrow
103	20	0.7	127	22	AAI22590	Probe #12523 for g
104	20	0.7	127	22	AAI47880	Probe #16566 used
105	20	0.7	127	22	AAI08268	Probe #8259 used t
106	20	0.7	256	22	AAK31061	Human diagnostic a
107	20	0.7	338	21	AAK00814	Human secreted pro
108	20	0.7	345	21	AAK01833	Human reproductive
109	20	0.7	355	21	AAH31025	Human colon cancer
110	20	0.7	370	21	AAK00812	Human secreted pro
111	20	0.7	408	22	AAK31074	Human diagnostic a
112	20	0.7	412	22	ABA09359	Human zn finger pr
113	20	0.7	413	22	AAK67544	Novel human polynu
114	20	0.7	425	22	AAI81543	Human polynucleoti
115	20	0.7	426	22	ABA4606	Human breast cell
116	20	0.7	426	22	ABA55057	Human foetal liver
117	20	0.7	426	22	ABA24812	Probe #3278 for ge
118	20	0.7	426	22	AAK03323	Human brain expres
119	20	0.7	426	22	AAK28774	Human bone marrow
120	20	0.7	426	22	AAI13359	Probe #3282 for ge
121	20	0.7	426	22	AAI34710	Probe #3396 used t
122	20	0.7	426	22	AAI03255	Probe #3246 used t
123	20	0.7	457	22	AAK42486	Human cDNA encodin
124	20	0.7	478	21	AAK00809	Human secreted pro
125	20	0.7	495	21	AAK00810	Human secreted pro
126	20	0.7	499	22	AAK02060	DNA encoding molec
127	20	0.7	518	22	ABA62830	Human foetal liver
128	20	0.7	524	22	AAH97864	Murine 7-transmemb
129	20	0.7	535	22	AAH97863	Murine 7-transmemb
130	20	0.7	570	22	AAK31073	Human diagnostic a
131	20	0.7	585	22	AAH06181	Human cDNA clone (
132	20	0.7	607	23	AAK88222	DNA encoding novel
133	20	0.7	659	22	AAK31068	Human diagnostic a
134	20	0.7	675	21	AAH31140	Human colon cancer
135	20	0.7	798	22	AAK42498	Human cDNA encodin
136	20	0.7	847	22	AAI20037	Human breast cancer
137	20	0.7	888	22	AAH72572	Human cervical can
138	20	0.7	1257	23	AAK86710	DNA encoding novel
139	20	0.7	1257	23	AAK92697	DNA encoding novel
140	20	0.7	1294	22	ABA08437	Human zn finger pr
141	20	0.7	1349	22	AAK03038	Human diagnostic a
142	20	0.7	1705	23	AAK84292	DNA encoding novel
143	20	0.7	1715	23	AAK92934	DNA encoding novel
144	20	0.7	1975	23	AAK87646	DNA encoding novel
145	20	0.7	2097	23	AAK87594	DNA encoding novel
146	20	0.7	2101	21	AAK92427	Human nucleic acid
147	20	0.7	2123	23	AAK92933	DNA encoding novel
148	20	0.7	2181	22	ABA82978	Human transcriptio
149	20	0.7	2228	22	AAK51909	Human polynucleoti
150	20	0.7	2235	23	AAK81692	DNA encoding novel
151	20	0.7	2274	23	AAK56578	DNA encoding novel
152	20	0.7	2313	23	AAK74278	DNA encoding novel
153	20	0.7	2313	23	AAK79107	DNA encoding novel
154	20	0.7	2313	23	AAK92935	DNA encoding novel
155	20	0.7	2345	22	AAH44772	Human DNA mismatch
156	20	0.7	2386	23	AAK86931	DNA encoding novel
157	20	0.7	2406	24	ABA95862	Human zinc finger
158	20	0.7	2478	22	AAH17764	Human cDNA sequenc
159	20	0.7	2865	22	AAK52893	Human polynucleoti
160	20	0.7	3082	22	AAK53018	Human polynucleoti
161	20	0.7	3342	23	AAK92930	DNA encoding novel
162	20	0.7	3727	22	AAK52034	Human polynucleoti
163	20	0.7	4272	22	AAK58698	Human polynucleoti
164	20	0.7	4420	22	AAK52954	Human polynucleoti
165	20	0.7	4422	22	AAK60484	Human polynucleoti
166	20	0.7	6135	22	ABA20822	Human nervous syst
167	20	0.7	6135	22	AAK29211	Genomic sequence #
168	20	0.7	6135	22	AAK85912	Human immune/haema
169	20	0.7	7470	22	AAK89416	Human digestive sy
170	20	0.7	7470	22	AAK89417	Human digestive sy
171	20	0.7	7470	22	AAK31921	Human liver associ
172	20	0.7	7470	22	AAK31922	Human liver associ
173	20	0.7	2858	22	AAK89418	Human digestive sy
174	20	0.7	2858	22	AAK31923	Human liver associ
175	19	0.7	69	22	AAK51303	DNA encoding zinc
176	19	0.7	127	22	AAK50423	Human bone marrow
177	19	0.7	223	21	AAK08609	Human secreted pro
178	19	0.7	275	20	AAK91335	Human polynucleoti
179	19	0.7	279	20	AAK214672	Human gene expres
180	19	0.7	290	21	AAK08608	Human secreted pro
181	19	0.7	392	21	AAK10140	Human secreted pro
182	19	0.7	427	23	AAK78109	DNA encoding novel
183	19	0.7	449	22	AAI92077	Human polynucleoti
184	19	0.7	451	22	AAI84583	Human polynucleoti
185	19	0.7	452	22	ABA58065	Human foetal liver
186	19	0.7	452	22	AAK06138	Human brain expres
187	19	0.7	452	22	AAK31792	Human bone marrow
188	19	0.7	452	22	AAI37667	Probe #6353 used t
189	19	0.7	479	22	AAK06869	Human brain expres
190	19	0.7	479	22	AAK32591	Human bone marrow
191	19	0.7	479	22	AAI38410	Probe #7096 used t
192	19	0.7	538	22	AAK37553	Human bone marrow
193	19	0.7	581	22	ABA64230	Human foetal liver
194	19	0.7	581	22	AAK12708	Human brain expres
195	19	0.7	581	22	AAI19222	Probe #9155 for ge
196	19	0.7	581	22	AAI44382	Probe #13068 used
197	19	0.7	597	22	AAI24099	Human breast cancer
198	19	0.7	599	22	AAK68309	Human lung tumour
199	19	0.7	633	22	AAK31056	Human diagnostic a
200	19	0.7	635	22	AAI15253	Human breast cancer
201	19	0.7	664	22	AAK26518	Human cDNA encodin
202	19	0.7	692	22	AAK26250	Human cDNA encodin
203	19	0.7	721	22	AAI95208	Human neuroblastom
204	19	0.7	753	22	AAH06677	Human cDNA clone (
205	19	0.7	792	22	AAH06456	Human cDNA clone (
206	19	0.7	833	22	AAI94504	Human neuroblastom
207	19	0.7	901	22	AAK31036	Human diagnostic a
208	19	0.7	957	22	AAK27636	DNA encoding human
209	19	0.7	976	22	AAK44904	Human conlig polyn
210	19	0.7	991	22	AAK31065	Human diagnostic a
211	19	0.7	1027	22	AAK31034	Human diagnostic a
212	19	0.7	1084	23	AAK68741	DNA encoding novel
213	19	0.7	1106	22	AAK44732	Human full-length
214	19	0.7	1265	22	ABA51279	Human breast cell
215	19	0.7	1265	22	ABA69287	Human foetal liver
216	19	0.7	1265	22	ABA36213	Probe #14679 for g
217	19	0.7	1265	22	AAK17574	Human liver expres
218	19	0.7	1265	22	AAK43387	Human bone marrow
219	19	0.7	1265	22	AAI24167	Probe #14100 for g
220	19	0.7	1365	22	AAI49454	Probe #18140 used
221	19	0.7	1365	22	AAI09731	Probe #9722 used t
222	19	0.7	1329	22	AAH33579	Human colon cancer
223	19	0.7	1390	22	AAH46940	Human secreted pro
224	19	0.7	1431	22	AAK52305	Human polynucleoti
225	19	0.7	1431	22	AAK53289	Human polynucleoti
226	19	0.7	1473	22	AAI58016	Human polynucleoti
227	19	0.7	1517	22	AAH45687	Human zinc finger
228	19	0.7	1561	24	AAK24005	Human inflammation

229	19	0.7	1564	22	AA164562	Human polypeptide-	c 302	19	0.7	6290	22	AA193812	Human polynucleoti
c 230	19	0.7	1591	24	AAD24004	Human inflammation	303	19	0.7	7574	22	AAK74485	Human immune/haema
c 231	19	0.7	1592	24	AAD23989	Human inflammation	304	19	0.7	7574	22	AAK74613	Human immune/haema
c 232	19	0.7	1604	22	ABA06533	Human CDNA SPO ID	c 305	19	0.7	8964	24	ABL33534	Human immune syste
c 233	19	0.7	1621	21	AAC58619	Human PRO1387 prot	c 306	19	0.7	8964	24	AAS61276	Human gene regulat
c 234	19	0.7	1630	21	AAAF7683	Human PRO1387 cDNA	c 307	19	0.7	9316	22	AAS28637	Genomic sequence #
c 235	19	0.7	1630	21	AAZ65110	Membrane-bound pro	c 308	19	0.7	10661	22	AAS28633	Genomic sequence #
c 236	19	0.7	1630	22	AAS21503	Human CDNA sequenc	c 309	19	0.7	12267	22	AAK85733	Human immune/haema
c 237	19	0.7	1630	22	AAF44256	Human PRO1387 (UNO	c 310	19	0.7	12275	22	AAK85734	Human immune/haema
c 238	19	0.7	1649	21	AAZ52456	HTFM clone 319415	c 311	19	0.7	13084	24	ABL34178	Human immune syste
c 239	19	0.7	1683	20	AAV73484	Human SRE-ZBP anal	c 312	19	0.7	20137	20	AAZ27027	Human CHD1 genomic
c 240	19	0.7	1751	21	AAAZ27133	Human inflammation	c 313	19	0.7	20138	20	AAZ26836	Human CHD1 Includi
c 241	19	0.7	1775	22	AAAH3508	DNA encoding GPCR	c 314	19	0.7	23071	20	AAZ27029	Human CHD1 genomic
c 242	19	0.7	1831	22	AAH15098	Human CDNA sequenc	c 315	19	0.7	26241	22	ABAI6222	Human nervous syst
c 243	19	0.7	1848	23	AAAF71552	DNA encoding novel	c 316	19	0.7	47090	22	AAK68125	Human immune/haema
c 244	19	0.7	1876	22	AA159802	Human polynucleoti	c 317	19	0.7	47090	22	AAK78219	Human immune/haema
c 245	19	0.7	1884	22	ABA82999	Human transcriptio	c 318	19	0.7	58708	22	AAK64739	Human immune/haema
c 246	19	0.7	1905	22	AAAF9495	Human hydrophobic	c 319	19	0.7	235033	19	AAV57926	Human immune/haema
c 247	19	0.7	1930	22	AAS26095	Human CDNA encodin	c 320	19	0.7	237326	19	AAV57903	Hereditary haemoch
c 248	19	0.7	1947	22	ABA46167	Human breast cell	c 321	19	0.7	2944528	24	ABA03041	Hereditary haemoch
c 249	19	0.7	1947	22	ABA56713	Human foetal liver	c 322	18	0.6	127	22	ABA47922	Listeria monocytog
c 250	19	0.7	1947	22	ABA26328	Probe #4794 for ge	c 323	18	0.6	127	22	ABA65815	Human foetal liver
c 251	19	0.7	1947	22	AAK04843	Human brain expres	c 324	18	0.6	127	22	ABA75221	Human foetal liver
c 252	19	0.7	1947	22	AAK30370	Human bone marrow	c 325	18	0.6	127	22	ABA32904	Probe #11370 for g
c 253	19	0.7	1947	22	AAI14979	Probe #4912 for ge	c 326	18	0.6	127	22	ABA39881	Probe #18347 for g
c 254	19	0.7	1947	22	AA136325	Probe #5011 used t	c 327	18	0.6	127	22	AAK14219	Human brain expres
c 255	19	0.7	2103	21	AAI04746	Probe #4737 used t	c 328	18	0.6	127	22	AAK23756	Human brain expres
c 256	19	0.7	2168	22	AAAC76302	Human ORF ORF1757	c 329	18	0.6	127	22	AAK39952	Human bone marrow
c 257	19	0.7	2180	22	AAV73485	Human SRE-ZBP anal	c 330	18	0.6	127	22	AAK49853	Human bone marrow
c 258	19	0.7	2180	22	AAH15638	Human CDNA sequenc	c 331	18	0.6	127	22	AAI20754	Probe #10687 for g
c 259	19	0.7	2222	22	AAI07016	Human reproductive	c 332	18	0.6	127	22	AAI26958	Probe #11691 for g
c 260	19	0.7	2222	22	AAI62707	Human breast or ov	c 333	18	0.6	127	22	AAI45977	Probe #14663 used
c 261	19	0.7	2223	22	AAAS89101	DNA encoding novel	c 334	18	0.6	127	22	AAI55772	Probe #24458 used
c 262	19	0.7	2226	22	AAAS52776	Human polynucleoti	c 335	18	0.6	127	22	AAI06455	Probe #6447 used t
c 263	19	0.7	2316	22	AAH16219	Human CDNA sequenc	c 336	18	0.6	235	16	AAI20602	Human gene signatu
c 264	19	0.7	2321	23	AAAS84294	DNA encoding novel	c 337	18	0.6	267	21	AAH12183	Human gene signatu
c 265	19	0.7	2342	22	AAI60594	Human polynucleoti	c 338	18	0.6	278	21	AAH74525	Human CDNA clone #
c 266	19	0.7	2352	22	AAK51792	Human polynucleoti	c 339	18	0.6	382	8	AAH71175	Sequence encoding
c 267	19	0.7	2359	22	AAI58808	Human polynucleoti	c 340	18	0.6	383	8	AAH70530	Sequence encoding
c 268	19	0.7	2416	22	AAI60592	Human polynucleoti	c 341	18	0.6	385	23	AAI68991	DNA encoding novel
c 269	19	0.7	2421	22	AAH16475	Human CDNA sequenc	c 342	18	0.6	450	22	AAI99892	Human protein enco
c 270	19	0.7	2426	22	AAI58806	Human polynucleoti	c 343	18	0.6	457	22	AAI94125	Human protein enco
c 271	19	0.7	2547	24	AAI99759	Mouse ischaemic co	c 344	18	0.6	472	21	AAI15320	Human protein enco
c 272	19	0.7	2660	23	AAI972617	DNA encoding novel	c 345	18	0.6	474	21	AAI80089	Human protein enco
c 273	19	0.7	2687	22	AAI58808	Human CDNA encodin	c 346	18	0.6	477	22	AAI58808	Human CDNA encodin
c 274	19	0.7	2710	23	AAI58808	DNA encoding novel	c 347	18	0.6	482	22	AAI58808	Human foetal liver
c 275	19	0.7	2779	23	AAI58808	DNA encoding novel	c 348	18	0.6	482	22	AAI58808	Human foetal liver
c 276	19	0.7	2925	22	AAI58808	Human polynucleoti	c 349	18	0.6	482	22	AAI58808	Human foetal liver
c 277	19	0.7	2936	24	AAI58808	Human CDNA encodin	c 350	18	0.6	482	22	AAI58808	Human foetal liver
c 278	19	0.7	2937	22	AAI60072	Human polynucleoti	c 351	18	0.6	482	22	AAI10236	Probe #169 for gen
c 279	19	0.7	3000	23	AAI60072	DNA encoding novel	c 352	18	0.6	482	22	AAI10236	Probe #170 used to
c 280	19	0.7	3046	22	AAI18332	Human CDNA sequenc	c 353	18	0.6	482	22	AAI10236	Probe #165 used to
c 281	19	0.7	3240	20	AAI26835	Human CHD1 transcr	c 354	18	0.6	489	22	ABA44091	Human breast cell
c 282	19	0.7	3244	20	AAI26833	Human CHD1 transcr	c 355	18	0.6	489	22	ABA44091	Human foetal liver
c 283	19	0.7	3264	20	AAI26834	Human CHD1 transcr	c 356	18	0.6	489	22	ABA44091	Human foetal liver
c 284	19	0.7	3268	20	AAI26832	Human CHD1 transcr	c 357	18	0.6	489	22	ABA44091	Human foetal liver
c 285	19	0.7	3351	22	AAH18183	Human CDNA sequenc	c 358	18	0.6	489	22	AAI28280	Human bone marrow
c 286	19	0.7	3352	22	AAH18183	DNA encoding novel	c 359	18	0.6	489	22	AAI28280	Probe #2778 for ge
c 287	19	0.7	3351	23	AAI586763	DNA encoding novel	c 360	18	0.6	489	22	AAI34201	Probe #2778 for ge
c 288	19	0.7	3354	23	AAI586763	DNA encoding novel	c 361	18	0.6	489	22	AAI34201	Probe #2778 for ge
c 289	19	0.7	3354	23	AAI586763	DNA encoding novel	c 362	18	0.6	489	22	AAI34201	Probe #2778 for ge
c 290	19	0.7	3614	22	AAH14654	Human CDNA sequenc	c 363	18	0.6	504	22	AAI34201	Probe #2778 for ge
c 291	19	0.7	3663	18	AAI72320	Embryonic stem cel	c 364	18	0.6	504	22	AAI34201	Probe #2778 for ge
c 292	19	0.7	3724	23	AAI05502	Drosophila melanog	c 365	18	0.6	510	22	ABA62707	Human foetal liver
c 293	19	0.7	3810	20	AAI05502	DNA encoding novel	c 366	18	0.6	510	22	ABA62707	Human foetal liver
c 294	19	0.7	4209	22	AAI05502	Human foetal liver	c 367	18	0.6	510	22	AAI05502	Human foetal liver
c 295	19	0.7	4338	22	AAI05502	Human foetal liver	c 368	18	0.6	510	22	AAI05502	Human foetal liver
c 296	19	0.7	4338	22	AAI05502	Human foetal liver	c 369	18	0.6	510	22	AAI05502	Human foetal liver
c 297	19	0.7	4830	23	AAI05502	Human foetal liver	c 370	18	0.6	510	22	AAI05502	Human foetal liver
c 298	19	0.7	4830	23	AAI05502	Human foetal liver	c 371	18	0.6	510	22	AAI05502	Human foetal liver
c 299	19	0.7	5421	22	AAI05502	DNA encoding novel	c 372	18	0.6	565	24	AAI05502	Human foetal liver
c 300	19	0.7	5421	22	AAI05502	DNA encoding novel	c 373	18	0.6	575	13	AAI05502	Human foetal liver
c 301	19	0.7	6287	22	ABA20058	Human nervous syst	c 374	18	0.6	577	22	AAI05502	Human foetal liver

375	18	0.6	577	22	AA526358	Human CDNA encodin	448	18	0.6	1785	22	AAH75047	Nucleotide sequenc
376	18	0.6	584	23	AA568996	DNA encoding novel	449	18	0.6	1785	22	AAH76370	Human PMSR3 CDNA.
377	18	0.6	594	22	AB611905	Human foetal liver	450	18	0.6	1810	21	AA44997	Arabidopsis thalia
378	18	0.6	594	22	ABA29442	Probe #7908 for ge	451	18	0.6	1817	22	AAK94651	Human full-length
379	18	0.6	594	22	AAK10216	Human brain expres	452	18	0.6	1818	23	ABL22397	Drosophila melanog
380	18	0.6	594	22	AAK36114	Human bone marrow	453	18	0.6	1865	23	AA568355	DNA encoding novel
381	18	0.6	594	22	AA117236	Probe #7169 for ge	454	18	0.6	1883	23	AA569005	DNA encoding novel
382	18	0.6	594	22	AA141830	Probe #10516 used	455	18	0.6	1997	21	AA572418	Human nucleic acid
383	18	0.6	614	22	AA525989	Human CDNA encodin	456	18	0.6	1999	22	AA526019	Human CDNA encodin
384	18	0.6	614	23	AA585429	DNA encoding novel	457	18	0.6	2000	20	AA220098	Rice serine palmit
385	18	0.6	620	23	AA570669	DNA encoding novel	458	18	0.6	2038	22	AA522799	Human CDNA encodin
386	18	0.6	638	22	AA526425	Human CDNA encodin	459	18	0.6	2133	18	AAV01890	Human OTK18 gene.
387	18	0.6	678	22	ABA08751	Human secreted pro	460	18	0.6	2156	9	AAH82302	Sequence of HaellI
388	18	0.6	738	20	AA598852	Human validated ca	461	18	0.6	2204	22	AA526025	Human CDNA encodin
389	18	0.6	738	22	AAK83140	Human immune/haema	462	18	0.6	2361	22	AAH14258	Human CDNA sequenc
390	18	0.6	738	22	AAK83141	Human immune/haema	463	18	0.6	2369	24	ABA01156	Human zinc finger
391	18	0.6	740	13	AA027236	DNA encoding human	464	18	0.6	2412	22	AA522563	Human CDNA encodin
392	18	0.6	748	20	AAK61450	DNA encoding a hum	465	18	0.6	2427	22	AA500460	Plasmid pLOI2223 u
393	18	0.6	781	21	AAK77651	Human cancer assoc	466	18	0.6	2647	21	AAA30827	Expression cassett
394	18	0.6	795	22	AAH75048	Nucleotide sequenc	467	18	0.6	2720	22	AAK72833	Secreted protein g
395	18	0.6	795	22	AAH76371	Human PMSR3 CDNA.	468	18	0.6	2852	22	AAK68501	Human immune/haema
396	18	0.6	795	23	AA564690	DNA encoding novel	469	18	0.6	2880	21	AAA23438	CDNA encoding huma
397	18	0.6	799	22	ABA49151	Human breast cell	470	18	0.6	2890	22	AAK93782	Human CDNA encodin
398	18	0.6	799	22	ABA67064	Human foetal liver	471	18	0.6	2913	22	AAK51782	Human CDNA encodin
399	18	0.6	799	22	ABA34157	Probe #12623 for g	472	18	0.6	2924	22	AAK72831	Human polynucleoti
400	18	0.6	799	22	AAK15508	Human brain expres	473	18	0.6	2926	22	AAK72805	Secreted protein g
401	18	0.6	799	22	AAK41238	Human bone marrow	474	18	0.6	2926	22	AAK72830	Secreted protein g
402	18	0.6	799	22	AA121990	Probe #11923 for g	475	18	0.6	2929	21	AAK59988	Human secreted pro
403	18	0.6	799	22	AA147280	Probe #15966 used	476	18	0.6	2950	21	AAK26427	Human secreted pro
404	18	0.6	799	22	AA107685	Probe #7676 used t	477	18	0.6	2950	22	AAK72834	Secreted protein g
405	18	0.6	898	22	AAK91659	Human CDNA 5'-end	478	18	0.6	2951	22	AAK72832	Secreted protein g
406	18	0.6	898	22	AAK93578	Human CDNA clone r	479	18	0.6	2964	22	AAK45127	Secreted protein g
407	18	0.6	908	21	AAK76458	Human ORFX ORF2013	480	18	0.6	3033	22	AAH89951	Human TANGO 273 cd
408	18	0.6	911	21	AAK98990	Human pancreatic c	481	18	0.6	3041	21	AAK96800	Human bone marrow
409	18	0.6	943	22	AA525907	Human CDNA encodin	482	18	0.6	3046	22	AAH16119	Nuclear transport
410	18	0.6	960	22	ABA834167	Human colon cancer	483	18	0.6	3060	21	AAK96801	Human CDNA sequenc
411	18	0.6	960	22	ABA83031	Human transcriptio	484	18	0.6	3246	22	AA500463	Nuclear transport
412	18	0.6	968	22	ABA44057	Human breast cell	485	18	0.6	3262	22	AA500463	Plasmid pLOI2226 u
413	18	0.6	968	22	ABA54512	Human foetal liver	486	18	0.6	3351	22	AAH23745	zinc finger protei
414	18	0.6	968	22	ABA24295	Probe #2761 for ge	487	18	0.6	3443	22	AA500464	Plasmid pLOI2227 u
415	18	0.6	968	22	AAK02803	Human brain expres	488	18	0.6	3508	22	AA500465	Plasmid pLOI2228 u
416	18	0.6	968	22	AAK28246	Human bone marrow	489	18	0.6	3524	21	AAK75757	Human ORFX ORF1312
417	18	0.6	968	22	AAI12811	Probe #2744 for ge	490	18	0.6	3699	23	ABL22396	Drosophila melanog
418	18	0.6	968	22	AAI34166	Probe #2852 used t	491	18	0.6	3721	23	AA584395	DNA encoding novel
419	18	0.6	968	22	AA102728	Probe #2719 used t	492	18	0.6	3754	18	AA585853	DNA encoding novel
420	18	0.6	985	22	AAK91182	ADAM gene #6. Hom	493	18	0.6	4549	21	AAV01891	Human OTK18 gene.
421	18	0.6	1037	21	AAZ55616	Human KIAA0001 rec	494	18	0.6	4549	21	AAK38981	thart and dhab1.2.
422	18	0.6	1037	21	AAZ55616	Human polynucleoti	495	18	0.6	4549	22	AAK82001	Vector pCCL1920 DNA
423	18	0.6	1055	21	AAK79870	Human secreted pro	496	18	0.6	4732	22	AAK09559	Human protocadheri
424	18	0.6	1087	22	AA526550	Human CDNA encodin	497	18	0.6	4732	22	AAK52766	Human polynucleoti
425	18	0.6	1322	22	ABA46662	Human breast cell	498	18	0.6	6511	22	AA528632	Genomic sequence #
426	18	0.6	1322	22	ABA64534	Human foetal liver	499	18	0.6	6511	22	AA528638	Genomic sequence #
427	18	0.6	1322	22	ABA31667	Probe #10133 for g	500	18	0.6	6775	22	AA199321	Human excretory re
428	18	0.6	1322	22	AAK12985	Human brain expres	501	18	0.6	6775	22	AA163671	Human kidney relat
429	18	0.6	1322	22	AAK38712	Human bone marrow	502	18	0.6	7011	23	AA568893	DNA encoding novel
430	18	0.6	1322	22	AAI19517	Probe #9450 for ge	503	18	0.6	7887	19	AAV27866	Mouse telomerase R
431	18	0.6	1322	22	AAI44708	Probe #13394 used	504	18	0.6	8503	22	AA502334	Plasmid pAN004 car
432	18	0.6	1322	22	AA105239	Probe #5230 used t	505	18	0.6	10200	22	AAK77727	Human immune/haema
433	18	0.6	1342	23	AA570666	DNA encoding novel	506	18	0.6	10325	22	AAK67805	Human immune/haema
434	18	0.6	1344	22	AA160398	Human polynucleoti	507	18	0.6	10801	22	AA502323	Plasmid pAN240 car
435	18	0.6	1369	23	AA584324	DNA encoding novel	508	18	0.6	12552	22	AAK80231	Human immune/haema
436	18	0.6	1468	23	AA585584	DNA encoding novel	509	18	0.6	12880	22	AA546771	Tumour suppressor
437	18	0.6	1468	23	AA585584	Human EST-derived	510	18	0.6	17310	22	AAK83881	Human immune/haema
438	18	0.6	1558	16	AAQ85986	Raphanus sativus p	511	18	0.6	21501	22	AAK87857	Human immune/haema
439	18	0.6	1599	22	AA545046	CDNA encoding nove	512	18	0.6	39887	22	AAK79153	Human immune/haema
440	18	0.6	1645	21	AAK47096	Arabidopsis thalia	513	18	0.6	49999	20	AAZ23869	Human LOBO homolog
441	18	0.6	1680	21	AAK48318	Arabidopsis thalia	514	18	0.6	109973	21	AAK22289	BAC containing rep
442	18	0.6	1683	21	AAK35472	Arabidopsis thalia	515	18	0.6	1230025	20	AAK22289	Nucleotide sequenc
443	18	0.6	1720	24	ABA05936	Human cell differe	516	18	0.6	1503900	22	AAK95240	Human neurogulin-1
444	18	0.6	1738	22	AAH15259	Human CDNA sequenc	517	18	0.6	1503900	22	AAK96733	Human neurogulin-1
445	18	0.6	1752	23	AA569003	DNA encoding novel	518	17	0.6	29	22	AAK02634	GSP1 (CG015R1) PCR
446	18	0.6	1769	23	AA584328	DNA encoding novel	519	17	0.6	45	22	AAK04690	Peptide B2.7 codin
447	18	0.6	1778	21	AAK59364	Human secreted pro	520	17	0.6	50	22	AAK33582	Human SNP oligonuc

521	17	0.6	53	21	AAA29670	Plasmid construct	c 594	17	0.6	596	21	AAE12728	Aspergillus oryzae
522	17	0.6	69	22	AA513071	DNA encoding zinc	595	17	0.6	596	22	AA526392	Human cDNA encodin
523	17	0.6	194	22	ABA75667	Human foetal liver	596	17	0.6	597	22	AA526251	PERV-MSM4 clone br
524	17	0.6	194	22	AAK34279	Human brain expres	597	17	0.6	598	20	AA597889	Human secreted pro
525	17	0.6	194	22	AAK50304	Human bone marrow	598	17	0.6	624	22	AA539408	cDNA encoding nove
526	17	0.6	194	22	AAI56271	Probe #24957 used	599	17	0.6	624	22	AAK88256	Human digestive sy
527	17	0.6	233	21	AAAC31623	Human secreted pro	600	17	0.6	632	22	AAH97876	Murine 7-transemb
528	17	0.6	257	21	AA4A2043	Human secreted exp	601	17	0.6	644	22	AA545029	cDNA encoding nove
529	17	0.6	289	19	AAV60049	Nucleic acid 645.	602	17	0.6	657	20	AA597842	Human secreted pro
530	17	0.6	297	21	AAA40426	Human secreted exp	603	17	0.6	657	22	AAH44439	Human alkylated-DN
531	17	0.6	314	21	AAAC00398	Human secreted pro	604	17	0.6	673	22	AAI36944	Human musculoskele
532	17	0.6	318	20	AA4A0476	Human secreted pro	605	17	0.6	676	22	AAE27650	DNA encoding human
533	17	0.6	324	22	AA544566	Human cDNA encodin	606	17	0.6	694	22	AA526254	Human cDNA encodin
534	17	0.6	343	22	AAI65415	Nucleotide sequenc	607	17	0.6	701	20	AA215844	Human gene express
535	17	0.6	347	20	AA515882	Human secreted pro	608	17	0.6	712	21	AA535908	Arabidopsis thalia
536	17	0.6	355	22	AA537836	Novel human diagno	609	17	0.6	714	22	AAK57598	Human immune/haema
537	17	0.6	373	23	AA567212	DNA encoding novel	610	17	0.6	715	22	AAK70696	Human immune/haema
538	17	0.6	401	22	ABAI2150	Human nervous syst	611	17	0.6	732	20	AAK25548	Rice abscisic acid
539	17	0.6	401	22	AAH99901	Human protein enco	612	17	0.6	733	21	AA526296	Human ORFX ORF1851
540	17	0.6	413	22	AAI92471	Human polynucleotl	613	17	0.6	739	22	AAH34145	Human colon cancer
541	17	0.6	417	20	AA591233	T. gondii DNA enco	614	17	0.6	740	22	AAH05067	Human cDNA clone (
542	17	0.6	417	22	AA542556	T. gondii DNA enco	615	17	0.6	742	22	AAH07538	Human cDNA clone (
543	17	0.6	441	22	ABA08283	Human Zn finger pr	616	17	0.6	759	22	AAH03899	Human cDNA clone (
544	17	0.6	448	22	ABA57646	Human foetal liver	617	17	0.6	803	22	AAI97420	Human neuroblastom
545	17	0.6	448	22	ABA27066	Probe #5532 for ge	618	17	0.6	836	22	AA526031	Human cDNA clone (
546	17	0.6	448	22	AAK05700	Human brain expres	619	17	0.6	841	22	AAH05482	Human cDNA clone (
547	17	0.6	448	22	AAK31317	Human bone marrow	620	17	0.6	847	22	AAH03433	Human cDNA clone (
548	17	0.6	448	22	AAI15632	Probe #5565 for ge	621	17	0.6	860	22	AA531058	Human diagnostic a
549	17	0.6	448	22	AAI37213	Probe #5899 used t	622	17	0.6	877	22	AAH06903	Human cDNA clone (
550	17	0.6	470	22	ABA26040	Probe #4506 for ge	623	17	0.6	885	23	AA578655	DNA encoding novel
551	17	0.6	470	22	AAK30084	Human bone marrow	624	17	0.6	889	20	AA574420	Rat U3 gene trap d
552	17	0.6	470	22	AAAD07626	Human secreted pro	625	17	0.6	905	22	AA526462	Human cDNA encodin
553	17	0.6	473	22	AAK31276	Human bone marrow	626	17	0.6	906	22	AA51548	Human breast cell
554	17	0.6	473	22	AAK63970	Human immune/haema	627	17	0.6	906	22	ABA59610	Human foetal liver
555	17	0.6	478	22	ABA52068	Human foetal liver	628	17	0.6	906	22	AAI24444	Probe #14377 for g
556	17	0.6	478	22	AAK21882	Probe #348 for gen	629	17	0.6	914	19	AAV29574	Phaffia rhodozyma
557	17	0.6	478	22	AAK00350	Human brain expres	630	17	0.6	930	21	AA548843	Arabidopsis thalia
558	17	0.6	478	22	AAK25794	Human bone marrow	631	17	0.6	938	21	AA575394	Human ORFX ORF949
559	17	0.6	478	22	AAI10423	Probe #356 for gen	632	17	0.6	953	22	AAH00777	Abstita corymbifer
560	17	0.6	478	22	AAI31675	Probe #361 used to	633	17	0.6	972	22	AA525928	Human cDNA encodin
561	17	0.6	478	22	AAI00358	Probe #349 used to	634	17	0.6	1001	22	ABA49927	Human breast cell
562	17	0.6	480	22	AAK29294	Human cDNA 3'-end	635	17	0.6	1001	22	ABA67846	Human foetal liver
563	17	0.6	505	22	ABA61306	Human foetal liver	636	17	0.6	1001	22	ABA634905	Human Ca channel s
564	17	0.6	505	22	AAK09603	Human brain expres	637	17	0.6	1019	22	AAK44366	Human secreted pro
565	17	0.6	505	22	AAK35496	Human bone marrow	638	17	0.6	1029	23	ABLI3627	Human cDNA encodin
566	17	0.6	505	22	AAI41210	Probe #9896 used t	639	17	0.6	1037	22	AAH52555	S. epidermidis ope
567	17	0.6	512	22	AAH09437	Human cDNA clone (640	17	0.6	1047	22	AA503034	Human diagnostic a
568	17	0.6	516	22	AAK60837	Human immune/haema	641	17	0.6	1065	21	AA545217	Arabidopsis thalia
569	17	0.6	518	22	AAH97872	Murine 7-transemb	642	17	0.6	1097	22	ABA08482	Human Ca channel s
570	17	0.6	526	21	AA539190	Zea mays DNA fragm	643	17	0.6	1103	21	AA569445	Human secreted pro
571	17	0.6	532	21	AAZ80564	Human colon cancer	644	17	0.6	1123	22	AA510875	Human cDNA encodin
572	17	0.6	541	21	AA569771	Human breast tumou	645	17	0.6	1124	22	ABA16976	Human nervous syst
573	17	0.6	544	22	AAK92455	Human cDNA 3'-end	646	17	0.6	1125	22	ABA16977	Human nervous syst
574	17	0.6	561	21	AAA69744	Human ovarian carc	647	17	0.6	1129	22	AA545217	cDNA encoding nove
575	17	0.6	563	21	AAAI6428	Human colon cancer	648	17	0.6	1140	22	AAH68001	S. epidermidis ope
576	17	0.6	566	22	ABA63219	Human foetal liver	649	17	0.6	1152	22	AAH53316	Arabidopsis thalia
577	17	0.6	566	22	AAK11684	Human brain expres	650	17	0.6	1155	21	AA549626	Human pancreatic c
578	17	0.6	566	22	AAK37427	Human bone marrow	651	17	0.6	1173	22	ABA07013	Human cDNA encodin
579	17	0.6	566	22	AAI43285	Probe #11972 used	652	17	0.6	1173	22	AA532373	Human cDNA encodin
580	17	0.6	567	22	ABA62385	Human foetal liver	653	17	0.6	1183	21	AA543780	Zea mays DNA fragm
581	17	0.6	567	22	ABA28699	Probe #8165 for ge	654	17	0.6	1202	21	AA574424	Arabidopsis thalia
582	17	0.6	567	22	AAK10694	Human brain expres	655	17	0.6	1219	22	AAE27657	DNA encoding human
583	17	0.6	567	22	AAK36581	Human bone marrow	656	17	0.6	1241	23	AA573410	Human immune/haema
584	17	0.6	567	22	AAI17433	Probe #7366 for ge	657	17	0.6	1294	22	AAK85176	Candida parapsillos
585	17	0.6	567	22	AAI42339	Probe #11025 used	658	17	0.6	1297	22	AAH00481	Human secreted pro
586	17	0.6	594	22	AAI35295	Human musculoskele	659	17	0.6	1300	20	AA510869	Human cDNA encodin
587	17	0.6	594	22	AAH10991	Human cDNA clone (660	17	0.6	1318	18	AA578761	Human AD4 gene gen
588	17	0.6	595	22	ABA63831	Human foetal liver	661	17	0.6	1320	20	AA578761	DNA encoding novel
589	17	0.6	595	22	ABA31012	Probe #9478 for ge	662	17	0.6	1329	21	AA578761	Micrococcus luteus
590	17	0.6	595	22	AAK12342	Human brain expres	663	17	0.6	1380	20	AA578761	Plasmidium falcipa
591	17	0.6	595	22	AAK38061	Human bone marrow	664	17	0.6	1383	22	AAI70124	Human reproductive
592	17	0.6	595	22	AAI18833	Probe #8766 for ge	665	17	0.6				
593	17	0.6	595	22	AAI43956	Probe #12642 used	666	17	0.6				

813	17	0.6	2506	22	AAH89919	Human bone marrow
814	17	0.6	2544	23	AAH82160	DNA encoding novel
815	17	0.6	2553	24	ABA05651	Human zinc finger
816	17	0.6	2560	22	AAK53091	Human polynucleoti
817	17	0.6	2561	22	AAH17818	Human cDNA sequenc
818	17	0.6	2565	21	AAZ96805	Nuclear transpor
819	17	0.6	2572	22	AAK53098	Human polynucleoti
820	17	0.6	2572	22	AAH81118	DNA encoding novel
821	17	0.6	2604	23	AAH92936	DNA encoding novel
822	17	0.6	2624	22	AAK52114	Human polynucleoti
823	17	0.6	2625	22	AAH44597	Human full-length
824	17	0.6	2644	22	AAH44769	Human conlig polyn
825	17	0.6	2680	19	AAV64579	MYC-binding zinc-f
826	17	0.6	2694	21	AAZ96806	Nuclear transpor
827	17	0.6	2718	23	AAH69395	DNA encoding novel
828	17	0.6	2741	21	AAZ52436	HTPM clone 1484257
829	17	0.6	2761	22	AAH60214	Human polynucleoti
830	17	0.6	2769	22	AAH58428	Human polynucleoti
831	17	0.6	2864	22	AAK52107	Human polynucleoti
832	17	0.6	2920	22	AAH85209	DNA encoding novel
833	17	0.6	2940	24	ABH99447	Mouse ischaemic co
834	17	0.6	2959	22	AAK51505	Human polynucleoti
835	17	0.6	2959	22	AAK52489	Human polynucleoti
836	17	0.6	2960	23	ABH15944	Drosophila melanog
837	17	0.6	2982	21	AAH72388	Human nucleic acid
838	17	0.6	3006	21	AAZ37101	CDNA encoding a hu
839	17	0.6	3027	18	AAH75305	Nucleotide sequenc
840	17	0.6	3041	12	AAH11852	Glutamate receptor
841	17	0.6	3044	18	AAH64551	Tomato S-ribonucle
842	17	0.6	3105	24	ABA03746	Human zinc finger
843	17	0.6	3150	21	AAH76991	Human ORF ORF2546
844	17	0.6	3154	22	ABA07338	Human pancreatic c
845	17	0.6	3154	22	AAH32765	Human genomic DNA
846	17	0.6	3166	22	AAH22919	DNA encoding novel
847	17	0.6	3175	22	AAH59723	Human polynucleoti
848	17	0.6	3189	23	AAH85210	DNA encoding novel
849	17	0.6	3197	22	AAH26673	Human genomic DNA
850	17	0.6	3202	22	AAH54022	S. epidermidis gen
851	17	0.6	3213	22	AAH57937	Human polynucleoti
852	17	0.6	3244	23	ABH08432	Drosophila melanog
853	17	0.6	3264	23	ABH06061	Drosophila melanog
854	17	0.6	3320	18	AAV09698	Porcine retrovirus
855	17	0.6	3331	22	AAH65067	DNA encoding novel
856	17	0.6	3369	22	AAH26673	Human breast cance
857	17	0.6	3387	22	AAH76384	Human T1LC polypep
858	17	0.6	3434	23	AAH73872	DNA encoding novel
859	17	0.6	3481	22	AAH72651	Human cervical can
860	17	0.6	3497	22	AAH55010	S. epidermidis gen
861	17	0.6	3545	23	ABH07616	Drosophila melanog
862	17	0.6	3600	20	AAH25180	HIV-1 group O isol
863	17	0.6	3639	23	AAH72011	DNA encoding novel
864	17	0.6	3651	22	AAH36943	Human musculoskele
865	17	0.6	3703	22	AAH44849	Human cDNA encodin
866	17	0.6	3726	23	AAH66984	DNA encoding novel
867	17	0.6	3840	24	ABH19881	Mouse ischaemic co
868	17	0.6	3871	23	ABH10528	Drosophila melanog
869	17	0.6	3900	23	AAH76383	Human T1LC polypep
870	17	0.6	3906	22	AAH73707	Human zinc finger
871	17	0.6	4008	22	ABH15890	Human nervous syst
872	17	0.6	4012	22	ABH07314	Human pancreatic c
873	17	0.6	4012	22	ABH07315	Human pancreatic c
874	17	0.6	4012	22	AAH90473	Human digestive sy
875	17	0.6	4012	22	AAH90474	Human digestive sy
876	17	0.6	4061	21	AAH49923	Human calcium chan
877	17	0.6	4086	23	AAH85765	DNA encoding novel
878	17	0.6	4156	22	AAH83412	Human immune/haema
879	17	0.6	4156	22	ABH18176	Drosophila melanog
880	17	0.6	4182	22	AAH26524	Human procon/oligo
881	17	0.6	4358	22	AAH14605	Human cDNA sequenc
882	17	0.6	4359	22	AAH05832	Human reproductive
883	17	0.6	4402	22	AAH67020	PEVY env protein c
884	17	0.6	4416	21	AAH09335	Human cancer assoc
885	17	0.6	4416	21	AAH09345	Human cancer assoc
886	17	0.6	4527	22	AAH58092	Human polynucleoti
887	17	0.6	4590	21	AAH75767	Human ORF ORF1322
888	17	0.6	4758	21	AAH69134	Human ABC1 gene ex
889	17	0.6	4918	22	AAH67022	PEVY env protein c
890	17	0.6	5167	22	AAH32679	Human genomic DNA
891	17	0.6	5167	22	AAH76525	Human immune/haema
892	17	0.6	5223	23	ABH13306	Drosophila melanog
893	17	0.6	5440	22	AAH05240	Human reproductive
894	17	0.6	5440	22	AAH05242	Human reproductive
895	17	0.6	5706	22	AAH76497	Human immune/haema
896	17	0.6	5770	23	AAH84301	DNA encoding novel
897	17	0.6	5980	24	ABH33188	Human immune syste
898	17	0.6	6070	22	AAH81491	Human immune syste
899	17	0.6	6076	22	AAH67021	PEVY env protein c
900	17	0.6	6076	22	AAH67032	PEVY env protein c
901	17	0.6	6173	22	AAH30504	DNA encoding novel
902	17	0.6	6173	22	AAH06284	Human reproductive
903	17	0.6	6173	22	AAH69343	Human immune/haema
904	17	0.6	6174	22	AAH30503	DNA encoding novel
905	17	0.6	6174	22	AAH06283	Human reproductive
906	17	0.6	6174	22	AAH69342	Human immune/haema
907	17	0.6	6393	22	AAH26744	Human genomic DNA
908	17	0.6	6426	22	AAH26754	Human genomic DNA
909	17	0.6	6427	22	AAH26755	Human genomic DNA
910	17	0.6	6689	22	ABH19618	Drosophila melanog
911	17	0.6	6815	22	AAH26745	Human genomic DNA
912	17	0.6	6815	22	AAH68521	Human immune/haema
913	17	0.6	6916	23	ABH19616	Drosophila melanog
914	17	0.6	6953	17	AAH17116	Rhodopsin gene. H
915	17	0.6	7120	23	ABH04480	Drosophila melanog
916	17	0.6	7149	23	ABH19614	Drosophila melanog
917	17	0.6	7207	21	AAH21316	Human low adenosin
918	17	0.6	7207	21	AAH35194	Human adenosine re
919	17	0.6	7297	22	AAH89535	Human digestive sy
920	17	0.6	7333	22	AAH77726	Defective retrovir
921	17	0.6	7362	22	AAH67019	PEVY env protein c
922	17	0.6	7393	18	AAH74883	Porcine retrovirus
923	17	0.6	7445	22	ABH07339	Human pancreatic c
924	17	0.6	7445	22	AAH32766	Human genomic DNA
925	17	0.6	7562	23	ABH18203	Drosophila melanog
926	17	0.6	7873	22	AAH67023	PEVY env protein c
927	17	0.6	8196	18	AAV09699	Porcine retrovirus
928	17	0.6	8209	18	AAV09700	Porcine retrovirus
929	17	0.6	8353	17	AAH26697	Human genomic DNA
930	17	0.6	8372	22	AAH46934	Tumour suppressor
931	17	0.6	8420	22	AAH89536	Human digestive sy
932	17	0.6	8446	24	ABH33670	Human immune syste
933	17	0.6	8456	24	ABH33977	Human immune syste
934	17	0.6	9078	23	ABH13626	Drosophila melanog
935	17	0.6	9183	23	ABH06060	Drosophila melanog
936	17	0.6	9437	23	ABH10030	Drosophila melanog
937	17	0.6	9519	22	AAH137169	Human musculoskele
938	17	0.6	9618	22	AAH26698	Human genomic DNA
939	17	0.6	9666	24	ABH33084	Human immune syste
940	17	0.6	9666	24	ABH34530	Human metastasis a
941	17	0.6	10205	23	ABH18202	Drosophila melanog
942	17	0.6	10726	23	ABH06616	Drosophila melanog
943	17	0.6	10782	22	AAH75423	Human immune/haema
944	17	0.6	11089	22	AAH79653	Human immune/haema
945	17	0.6	11340	22	AAH75424	Human immune/haema
946	17	0.6	11754	22	AAH04033	Human ABC1 gene, p
947	17	0.6	12221	22	AAH41687	Genomic sequence #
948	17	0.6	12221	22	AAH81809	Human immune/haema
949	17	0.6	12221	22	AAH84385	Human immune/haema
950	17	0.6	12584	24	AAH61399	Human gene regulat
951	17	0.6	12586	22	AAH67076	Human immune/haema
952	17	0.6	12806	22	ABH24266	Drosophila melanog
953	17	0.6	13160	22	ABH15804	Human nervous syst
954	17	0.6	13160	22	ABH15805	Human nervous syst
955	17	0.6	13418	22	AAH86473	Human immune/haema
956	17	0.6	13526	22	AAH91515	Human digestive sy
957	17	0.6	13526	22	AAH32169	Human liver associ
958	17	0.6	14332	22	AAH41688	Genomic sequence #

959	17	0.6	14332	22	AAK81810	Human immune/haema
960	17	0.6	14332	22	AAK84386	Human immune/haema
961	17	0.6	14337	22	AAK91514	Human digestive sy
962	17	0.6	14337	22	AAK32168	Human liver associ
963	17	0.6	15254	22	AAK83135	Human immune/haema
964	17	0.6	15254	22	AAK83137	Human immune/haema
965	17	0.6	15255	22	AAK83136	Human immune/haema
966	17	0.6	15256	22	AAK83134	Human immune/haema
967	17	0.6	15418	21	AAK63785	Nucleotide sequenc
968	17	0.6	15630	21	AAK21317	Human low adenosin
969	17	0.6	15630	21	AAK35195	Human adenosine re
970	17	0.6	16271	22	ABA19209	Human nervous syst
971	17	0.6	16271	22	ABA19210	Human nervous syst
972	17	0.6	18663	22	AAK57589	Human immune/haema
973	17	0.6	18996	22	AAK57112	DNA encoding Droso
974	17	0.6	18996	23	ABL11802	Drosophila melanog
975	17	0.6	20811	23	ABL15604	Drosophila melanog
976	17	0.6	21509	23	ABL08688	Human reproductive
977	17	0.6	22635	22	AAH07203	Human reproductive
978	17	0.6	26040	22	AAH27887	Nucleotide sequenc
979	17	0.6	26555	22	AAK68372	Human immune/haema
980	17	0.6	26555	22	AAK68605	Human immune/haema
981	17	0.6	26555	22	AAI62833	Human genomic DNA
982	17	0.6	32194	22	AAH04340	Human reproductive
983	17	0.6	33923	22	AAK67071	Human immune/haema
984	17	0.6	43804	18	AAK76375	Chicken embryo let
985	17	0.6	43804	20	AAK26690	Complete genome se
986	17	0.6	44018	22	AAK82392	Avian adenovirus C
987	17	0.6	49136	21	AAK27475	NIDDM1 region incl
988	17	0.6	56632	22	AAK65581	Human immune/haema
989	17	0.6	72928	20	AAI18355	Human ASTRIJ 5' ge
990	17	0.6	72928	21	AAK80253	Human ASTRIJ 5' ge
991	17	0.6	80578	22	AAH44800	Human GPCR protein
992	17	0.6	81001	22	AAK30035	Human apolipoprote
993	17	0.6	90050	21	AAK91925	Wild type (C57BL/6
994	17	0.6	160271	22	AAK85750	Bipolar affective
995	17	0.6	160271	22	AAK85756	Human chromosome 1
996	17	0.6	160271	22	AAK504858	Human chromosome 1
997	17	0.6	160271	22	AAK504864	Human chromosome 1
998	17	0.6	160271	22	AAK506667	Human chromosome 1
999	17	0.6	160271	22	AAH40997	160kb fragment of
1000	17	0.6	160271	22	AAH23764	Human chromosome 1

ALIGNMENTS

RESULT 1

ID AAS74823 standard; cDNA; 2614 BP.

AC AAS74823;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #10627.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSF-) HYSFO INC.

XX

PI Drmanac RT, Liu C, Tang YT:
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG10636.
 XX
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 1; SEQ ID No 10627; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 2614 BP; 745 A; 611 C; 660 G; 598 T; 0 other;

Query Match 56.9%; Score 1577; DB 23; Length 2614;

Best local Similarity 99.9%; Pred. No. 0;

Matches 1817; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY	427	cagaatcgaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	486
DB	229		288
QY	487	aagctcgaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	546
DB	289		348
QY	547	cagg-aaatcctcctcgaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	605
DB	349		408
QY	606	ccattctgctttagtgcaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	665
DB	409		468
QY	666	tttgaggagtgcaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	725
DB	469		528
QY	726	cagcagcagcagtcgaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	785
DB	529		588
QY	786	agggcagcagtcgaagcgaagaattcaacttagtcctcctcgtccctcgtatttctcgaatcagc	845
DB	589		648
QY	846	gaatcaaatatgaagagtttgaggcagcgttatacgaagcgaagaattcgaatcagc	905
DB	649		708

QY	906	cagaagacacaaactvgvgagaaacttcatcgttacactgagtvgvgagaaagccttvgc	965
Dp	709	cagaagacacaaactvgvgagaaacttcatcgttacactgagtvgvgagaaagccttvgc	768
OY	966	agatgacgttcctcatcacaaccccaagagaaactctcggvggaaagccttatgtgvc	1025
Dp	769	agatgacgttcctcatcacaaccccaagagaaactctcggvggaaagccttatgtgvc	828
QY	1026	agggaaatgvtgvgcvgagcctttagtvtgaaagctaaacctgatcacatctaaagagaaac	1085
Dp	829	agggaaatgvtgvgcvgagcctttagtvtgaaagctaaacctgatcacatctaaagagaaac	888
QY	1086	tcagvgvgagaaaccttatgtgtgcaagagatvtgtgacagagccttacttvgaaagtcgaac	1145
Dp	889	tcagvgvgagaaaccttatgtgtgcaagagatvtgtgacagagccttacttvgaaagtcgaac	948
QY	1146	ctctttacaaactcgaagaaacactcgaaggctcgaagccttatgtvtgcaaggaatgtvgg	1205
Dp	949	ctctttacaaactcgaagaaacactcgaaggctcgaagccttatgtvtgcaaggaatgtvgg	1008
QY	1206	cagaagccttagcctgaaagtccaacacctcatatccacacagagvgcgacacatvtgvgaaag	1265
Dp	1009	cagaagccttagcctgaaagtccaacacctcatatccacacagagvgcgacacatvtgvgaaag	1068
QY	1266	ccttatgttvtgcaagggaaatgtvtgvgcgttgcgcgaagcalttcaacctgttcagaaac	1325
Dp	1069	ccttatgttvtgcaagggaaatgtvtgvgcgttgcgcgaagcalttcaacctgttcagaaac	1128
QY	1326	aagvgagacacttcaagaaagaaagccttcaacttgaaggaatgtgtgagcaagcctttagc	1385
Dp	1129	aagvgagacacttcaagaaagaaagccttcaacttgaaggaatgtgtgagcaagcctttagc	1188
QY	1386	cagaagctcaacacctcatcaaacactctaaagaaacacacacggvgagaaagccttatgtagc	1445
Dp	1189	cagaagctcaacacctcatcaaacactctaaagaaacacacacggvgagaaagccttatgtagc	1248
QY	1446	acagaatvtgvgcgttcaactttagcttgaagaaatcaaaccttcaaacacacacagagagaaac	1505
Dp	1249	acagaatvtgvgcgttcaactttagcttgaagaaatcaaaccttcaaacacacacagagagaaac	1308
QY	1506	tcagvgggtttaaacttatgtctcgtcgtvgaaagtvcgvgcagvtgctttagcctgaagtcgaac	1565
Dp	1309	tcagvgggtttaaacttatgtctcgtcgtcgtvgaaagtvcgvgcagvtgctttagcctgaagtcgaac	1368
QY	1566	cttaacaaacaccc-aaagvggtcacacacacggvggaaagcatttgtatgtgaacgaatgtgvg	1624
Dp	1369	cttaacaaacaccc-aaagvggtcacacacacggvggaaagcatttgtatgtgaacgaatgtgvg	1428
QY	1625	gcgaagcctttaaaccggagaaatcaaacaccttgatcaagcacacacactcagvggagaa	1684
Dp	1429	gcgaagcctttaaaccggagaaatcaaacaccttgatcaagcacacacactcagvggagaa	1488
QY	1685	gcacattgtatgtgctvgagtgtgtgacagagcctttaaagtcaagtcacaccttcatcaca	1744
Dp	1489	gcacattgtatgtgctvgagtgtgtgacagagcctttaaagtcaagtcacaccttcatcaca	1548
QY	1745	ccgaagagacacattcagaagggaaaagcctttaaagtvtgcaagvggtvtgvgaaagagtttgcg	1804
Dp	1549	ccgaagagacacattcagaagggaaaagcctttaaagtvtgcaagvggtvtgvgaaagagtttgcg	1608
QY	1805	gcagaagacctaaccgttvtlttvgacagaagvggcacactcagaatgtccttvtgtvgagaga	1864
Dp	1609	gcagaagacctaaccgttvtlttvgacagaagvggcacactcagaatgtccttvtgtvgagaga	1686
QY	1865	gtgtgtggtcagaagccttvtgtctaatcatctcatataaacacagagagacacgcaag	1924
Dp	1669	gtgtgtggtcagaagccttvtgtctaatcatctcatataaacacagagagacacgcaag	1728
QY	1925	ggggaaagccttaatvtgtgacagggagvtgtvggcaagcctttagcggcaatcacaacctat	1984
Dp	1729	ggggaaagccttaatvtgtgacagggagvtgtvggcaagcctttagcggcaatcacaacctat	1788
QY	1985	tagacacacagagacacattcagaagagaaagccttlatlttcagaabaagtvtgvgacggg	2044

[illegible]

```
QY 2198 ctgagataccagaagtgagacattctgtgtgattatgcatgagactgtactgtgaaga 2257
DB 129 gctagataccagaagtgagacattctgtgtgattatgcatgagactgtactgtgaaga 188
QY 2258 cttgtatctcatccacactgaagaagaattgtctgtctcatttcagaagccctgtccctc 2317
DB 189 cttgtatctcatccacactgaagaagaattgtctgtctcatttcagaagccctgtccctc 248
QY 2318 ctgacgtgtgattgtgtgtgtgtgaaccggtcaggtatgtatgtcagagagagcagt 2377
DB 249 ctgacgtgtgattgtgtgtgtgtgaaccggtcaggtatgtatgtcagagagagcagt 308
QY 2378 caaatgccagcagagataggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2437
DB 309 caaatgccagcagagataggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 368
QY 2438 gtccggcgttaaatccctacatactgaattgagaacctgtctccattgtgtgtcttct 2497
DB 369 gtccggcgttaaatccctacatactgaattgagaacctgtctccattgtgtgtcttct 428
QY 2498 cctgattgaccccaacctcactatcttaagatatacctgccttcctcctaattgtttt 2557
DB 429 cctgattgaccccaacctcactatcttaagatatacctgccttcctcctaattgtttt 488
QY 2558 acactgtgtgtccacacttttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2617
DB 489 acactgtgtgtccacacttttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 548
QY 2618 attccctattctgagccataaagaacca 2648
DB 549 attccctattctgagccataaagaacca 579

RESULT 3
AAS68256
ID AAS68256 standard; cDNA; 653 BP.
XX
AC AAS68256;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4060.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KN food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEO INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG04069.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
PS Claim 1; SEQ ID No 4060; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
```

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX
SQ Sequence 653 BP; 192 A; 151 C; 176 G; 134 T; 0 other;

Query Match 18.9%; Score 525; DB 23; Length 653;
Best Local Similarity 99.8%; Pred. No. 3; 3e-248;
Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1004 tgggggaaagccttatgtgtgtcagggaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1063
DB 52 tgggggaaagccttatgtgtgtcagggaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 111
QY 1064 gatcacacatcagagacacactcagggagaaaccttattgtgtcaggaattgtgtgtgtgt 1123
DB 112 gatcacacatcagagacacactcagggagaaaccttattgtgtcaggaattgtgtgtgtgt 171
QY 1124 agccttacttggaaatgcgaacctcttcaacatcagcggaacactcagggctcaagcc 1183
DB 172 agccttacttggaaatgcgaacctcttcaacatcagcggaacactcagggctcaagcc 231
QY 1184 ttatgtgtcaggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1243
DB 232 ttatgtgtcaggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 291
QY 1244 gaggggcacactcagggagaaaccttatgtgtcagggaaatgtgtgtgtgtgtgtgtgt 1303
DB 292 gaggggcacactcagggagaaaccttatgtgtcagggaaatgtgtgtgtgtgtgtgtgt 351
QY 1304 gcatcacacctgtgtcagacacagaagacacattcagagagaagaccttattgtcag 1363
DB 352 gcatcacacctgtgtcagacacagaagacacattcagagagaagaccttattgtcag 411
QY 1364 ggaagtgtgagcaggtttagccagaagtacacctcattcaaaccttaagaacacac 1423
DB 412 ggaagtgtgagcaggtttagccagaagtacacctcattcaaaccttaagaacacac 471
QY 1424 aggagagaagccttatgtatgcagagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1483
DB 472 aggagagaagccttatgtatgcagagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 531
QY 1484 caaaacacacacagagacacactcaggggttaaacttattgtctgtcgtgtgtgtgtgt 1543
DB 532 caaaacacacacagagacacactcaggggttaaacttattgtctgtcgtgtgtgtgtgt 591
QY 1544 gtgctttagcctgaagtcgaaccttaacaacacca 1579
DB 592 gtgctttagcctgaagtcgaaccttaacaacacca 627

RESULT 4
AAS26286
ID AAS26286 standard; cDNA; 983 BP.
XX
```

AC AAS26286;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 465.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnery; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241189.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

DR P-PSDB; AAU16299.

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

PS Claim 1; SEQ ID NO 465; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their
XX encoded secreted proteins. The nucleic acids and proteins are used to
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. They
XX are also used in diagnosing a pathological condition or susceptibility
XX to a pathological condition. Antibodies to the proteins can also
XX be used in alleviating symptoms associated with the disorders and in
XX diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
XX immunosorbent assays (ELISA). Disorders which are diagnosed or treated
XX include autoimmune diseases e.g. Rheumatoid arthritis,
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi
XX and ocular disorders e.g. corneal infection, and many other
XX disorders listed in the specification. The polypeptides can also
XX be used to aid wound healing and epithelial cell proliferation, to
XX prevent skin aging due to sunburn, to maintain organs before
XX transplantation, for supporting cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage
XX capabilities, fat content, lipid, protein, carbohydrate, vitamins,
XX minerals, cofactors and other nutritional components. The present
XX sequence encodes a novel secreted protein of the invention.

Query Match 8.3%; Score 231; DB 22; Length 983;

Best Local Similarity 99.6%; Pred. No. 2.4e-103; Mismatches 0; Gaps 0;

Matches 281; Conservative 0; Indels 103;

OY 287 gaacggaggagtgatgctgagactataacatctggtctcaactggaatccatctc 346
DB 570 gaacggaggagtgatgctgagactataacatctggtctcaactggaatccatctc 629
OY 347 taaccnaaaactatgtctcagctgagcgagggaaagccctctgagagagagaa 406
DB 630 taaccnaaaactatgtctcagctgagcgagggaaagccctctgagagagagaa 689
OY 407 atgtccactgagccctctgctcgaatcgagccagaatcaactagtcctctgccc 466
DB 690 atgtccactgagccctctgctcgaatcgagccagaatcaactagtcctctgccc 749
OY 467 tctgatttctcctcagcagcctcctcagccaacatgtgtgctgagtcacatctcca 526
DB 750 tctgatttctcctcagcagcctcctcagccaacatgtgtgctgagtcacatctcca 809

OY 527 gctgtttcaagttataggcaggaatcctctccaccctgg 568
DB 810 gctgtttcaagttataggcaggaatcctctccaccctgg 851

RESULT 5

AAI83890 standard; cDNA; 394 BP.

XX AAI83890;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 3950.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR P-PSDB; AA003959.

PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

PS Claim 1; SEQ ID NO 3950; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from Wipo
XX at ftp.wipo.int/pub/published_pct_sequences.

Sequence 394 BP; 148 A; 75 C; 93 G; 78 T; 0 other;

Query Match 3.6%; Score 99; DB 22; Length 394;

Best Local Similarity 99.3%; Pred. No. 2.6e-38; Mismatches 1; Indels 0; Gaps 0;

Matches 149; Conservative 0; Indels 1;

OY 2619 ttcacctatctgagcccaataaaagaccagactcagctgagtgagagaaatcacc 2678
DB 1 ttcacctatctgagcccaataaaagaccagactcagctgagtgagagaaatcacc 60
OY 2679 ctgctgtgagggttgggagccctcctcgtgacccctcctccactgaagcttctttt 2738
DB 61 ctgctgtgagggttgggagccctcctcgtgacccctcctccactgaagcttctttt 120

OY 2739 ccaataaattctttcaccatccca 2768
|
DB 121 cccaataaattctttcaccatccca 150

RESULT 6
AAS13031

ID AAS13031 standard: DNA; 69 BP.

AC AAS13031;

DT 17-DEC-2001 (first entry)

DE DNA encoding zinc finger domain TG-ZFD-012.

KM Zinc finger domain; cancer; human; ds; TG-ZFD-012.

OS Homo sapiens.

PN W0200160970-A2.

PD 23-AUG-2001.

PF 17-FEB-2001; 2001WO-KR00244.

PR 18-FEB-2000; 2000KR-0007730.

PA (TOOL-) TOOLGEN INC.

PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;

DR WPI: 2001-557644/62.

DR P-PSDB; AAU08514.

PT Identifying a zinc finger domain for e.g. designing new polypeptides
that bind to a specific site on a DNA, comprises expressing hybrid
nucleic acids with a test zinc finger domain in cells -

PS Example 22; Page 55; 147pp; English.

CC The invention relates to a method of identifying a zinc finger domain
CC that recognises a target site on a DNA. The method comprises expressing
CC hybrid nucleic acids with a test zinc finger domain in cells containing a
CC reporter construct, where the reporter gene is expressed above a given
CC level when a transcription factor recognises a recruitment and a target
CC site of a promoter, and not only the recruitment site of the promoter.
CC The method is used to: (a) identify a zinc finger domain that recognises
CC a target site on a DNA; (b) determine whether a test zinc finger domain
CC recognises a target site on a promoter; (c) generate a nucleic acid that
CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC recognised by zinc finger domains. The method can be used to design novel
CC polypeptides that bind to a specific site on a DNA. The method can
CC facilitate the customised generation of new polypeptides that can
CC regulate the expression of a selected target e.g. a gene required by a
CC pathogen can be repressed, a gene required for cancerous growth can be
CC repressed, or a gene poorly expressed or encoding a mutated protein can
CC be activated and overexpressed. The method can be used in vivo which
CC enables identification of polypeptides that bind to a specific site on a
CC DNA in the intracellular milieu. The present sequence represents the
CC coding sequence of TG-ZFD-012 zinc finger protein which was used in
CC the method of the invention.

CC Sequence 69 BP; 17 A; 16 C; 20 G; 16 T; 0 other;

Query Match 2.5%; Score 69; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1269 tatgttcaggaagtgtgagcttcgcacacatcagctgcagacaag 1328
|
DB 1 tatgttcaggaagtgtgagcttcgcacacatcagctgcagacaag 60

OY 1329 aggaacacat 1337
|
DB 61 aggaacacat 69

RESULT 7
AAS13088

ID AAS13088 standard: DNA; 69 BP.

AC AAS13088;

DT 17-DEC-2001 (first entry)

DE DNA encoding zinc finger domain TG-ZFD-046.

KM Zinc finger domain; cancer; human; ds.

OS Homo sapiens.

PN W0200160970-A2.

PD 23-AUG-2001.

PF 17-FEB-2001; 2001WO-KR00244.

PR 18-FEB-2000; 2000KR-0007730.

PA (TOOL-) TOOLGEN INC.

PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;

DR WPI: 2001-557644/62.

DR P-PSDB; AAU08558.

PT Identifying a zinc finger domain for e.g. designing new polypeptides
that bind to a specific site on a DNA, comprises expressing hybrid
nucleic acids with a test zinc finger domain in cells -

PS Example 54; Page 69; 147pp; English.

CC The invention relates to a method of identifying a zinc finger domain
CC that recognises a target site on a DNA. The method comprises expressing
CC hybrid nucleic acids with a test zinc finger domain in cells containing a
CC reporter construct, where the reporter gene is expressed above a given
CC level when a transcription factor recognises a recruitment and a target
CC site of a promoter, and not only the recruitment site of the promoter.
CC The method is used to: (a) identify a zinc finger domain that recognises
CC a target site on a DNA; (b) determine whether a test zinc finger domain
CC recognises a target site on a promoter; (c) generate a nucleic acid that
CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC recognised by zinc finger domains. The method can be used to design novel
CC polypeptides that bind to a specific site on a DNA. The method can
CC facilitate the customised generation of new polypeptides that can
CC regulate the expression of a selected target e.g. a gene required by a
CC pathogen can be repressed, a gene required for cancerous growth can be
CC repressed, or a gene poorly expressed or encoding a mutated protein can
CC be activated and overexpressed. The method can be used in vivo which
CC enables identification of polypeptides that bind to a specific site on a
CC DNA in the intracellular milieu. The present sequence represents the
CC coding sequence of TG-ZFD-046 zinc finger protein which was used in
CC the method of the invention.

CC Sequence 69 BP; 21 A; 14 C; 18 G; 16 T; 0 other;

Query Match 2.5%; Score 69; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 tatatttcaggaagtgtgagcttcagtcggaagtcacacattacagacatcaag 2078
|
DB 1 tatatttcaggaagtgtgagcttcagtcggaagtcacacattacagacatcaag 60

QY 2079 aggcacac 2087
 |||||||
 Db 61 aggcacac 69

RESULT 8
 AAS68257
 ID AAS68257 standard; cDNA; 821 BP.
 XX
 AC AAS68257;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #4061.
 XX
 KW Human: Chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB; ABG04070.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS
 PS Claim 1; SEQ ID No 4061; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging or sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94364 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 821 BP; 174 A; 234 C; 215 G; 198 T; 0 other;

Query Match 1.8%; Score 49; DB 23; Length 821;
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 acacagccagcacagtcaccaagggaacacacacagctggtgatataagg 770
 |||||||
 Db 188 acacagccagcacagtcaccaagggaacacacacagctggtgatataagg 236

RESULT 9
 AAH33832
 ID AAH33832 standard; cDNA; 800 BP.
 XX
 AC AAH33832;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:888.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI: 2001-235357/24.
 DR P-PSDB; AAG74401.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PT
 PS
 PS Claim 1; Page 2813-2814; 9803bp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated cell,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAH77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 800 BP; 240 A; 173 C; 219 G; 166 T; 2 other;

Query Match 1.0%; Score 29; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 0.00079;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1659 cacacagagcacacactcaggaggaagcc 1687
 |||||||
 Db 413 cacacagagcacacactcaggaggaagcc 441

RESULT 10
ID AAS90429
XX AAS90429 standard; cDNA; 2349 BP.
AC AAS90429;
XX
XX 13-FEB-2002 (first entry)
DT
XX
DE DNA encoding novel human diagnostic protein #26233.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-0S08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR P-PSDB: ABG26242.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID NO 26233; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct_sequences.
XX
SQ Sequence 2349 BP; 664 A; 547 C; 608 G; 530 T; 0 other;

Query Match 1.0%; Score 29; DB 23; Length 2349;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagacacactcgaaggagagc 1687
|||||
DB 1645 caccagagacacactcgaaggagagc 1673

RESULT 11

AA158664
XX AA158664 standard; cDNA; 3582 BP.
XX
AC AA158664;
XX
XX 22-OCT-2001 (first entry)
DT
XX
DE Human polynucleotide SEQ ID NO 867.
XX
XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR P-PSDB: AAM39508.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
XX Claim 1; SEQ ID NO 867; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 3582 BP; 1024 A; 826 C; 910 G; 821 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagacacactcgaaggagagc 1687
|||||

DB 1487 caccagagacacactcagggaagcc 1515

RESULT 12

AAI60450/C

ID AAI60450 standard; cDNA; 3582 BP.

XX

AC AAI60450;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 4439.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW Leukemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000MO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR P-PSDB; AAM41294.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

PS Claim 1; SEQ ID NO 4439; 10078bp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

SQ Sequence 3582 BP; 821 A; 908 C; 826 G; 1026 T; 1 other;

XX

Query Match 1.0%; Score 29; DB 22; Length 3582;

Best Local Similarity 100.0%; Pred. No. 0.00078;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcagggaagcc 1687

DB 2096 CACCAGAGACACTCAGGGAAGCC 2068

RESULT 13

AAL00216

ID AAL00216 standard; cDNA; 159 BP.

XX

AC AAL00216;

XX

DT 21-NOV-2001 (first entry)

XX

DE Human reproductive system related antigen cDNA SEQ ID NO: 217.

XX

KW Human; reproductive system related antigen; reproductive system disorder;

KW cancer; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN WO200155320-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001MO-US01339.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251031.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465570/50.
 DR P-PSDB; AAM94246.
 DR
 XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 PS
 XX
 XX
 CC Claim 1; SEQ ID NO 217; 1297pp + Sequence Listing; English.
 CC
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a coding sequence of the
 CC invention.
 CC
 CC
 SQ Sequence 159 BP; 47 A; 33 C; 47 G; 30 T; 2 other;
 OY
 Db 1659 caccagagagacactcaaggaggaagc 1686
 18 caccagagagacactcaaggaggaagc 45
 RESULT 14
 AAS25949
 ID AAS25949 standard; cDNA; 1183 BP.
 XX
 AC AAS25949;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human cDNA encoding a novel secreted protein, Seq ID 128.
 XX
 KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
 KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnery; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischemia; anglogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.
 OS Homo sapiens.
 XX
 XX
 PN WO200155322-A2.
 XX

XX	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0238935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0256768.
PA	(HUMA-) HUMAN GENOME SCL INC.	
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
XX	WPI: 2001-488783/53.	
XX		
XX	New nucleic acid molecules encoding 461 human secreted proteins for	
PT	diagnosing, preventing, treating or ameliorating medical conditions and	
PT	used as food additives or preservatives -	
XX		
XX	Disclosure; SEQ ID No 1609; 980bp; English.	

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence is a genomic DNA encoding a partial novel secreted protein of the invention.

Query Match	1.0%;	Score 27;	DB 22;	Length 2717;
Best Local Similarity	100.0%;	Pred. No. 0.0076;		
Matches	27;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

Qy	2149	tcattcacaccagagagacacacag	2175
Db	962	tcattcacaccagagagacacacag	988

RESULT	16
AAV64361	.
ID	AAV64361 standard; cDNA; 3099 BP.

AAV64361; AC

DT 15-FEB-1999 (first entry)

DE Human stem cell zinc finger SZF1-2 CDNA.

KW SZF1-2; human; stem cell zinc finger; transcription factor;

KW haematopoiesis; CD34+; gene therapy; ss.

OS Homo sapiens.

...		
FH	Key	Location/Qualifiers
FT	CDS	379..1465
FT		/*tag= a

PN W09845326-A1.

PD 15-OCT-1998.

PF 08-APR-1998; 98WO-US06925.

PR 08-APR-1997: 97US-0041811.

PA (OSIR-) OSIRIS THERAPEUTICS INC

PI
civilian CI, Liu C, Small D

WPI: 1999-009328/01.

XX
DK

PT	New polynucleotides encoding stem cell zinc fingers - useful for
PT	gene therapy to replace or supplement defective gene for those
PT	proteins

PS Disclosure: Page 44-46; 76pp; English.

XX This cDNA sequence includes a coding region for human stem cell
CC zinc finger protein SZF1-2 (see AAW81632), a novel transcription
CC factor that appears to be expressed in most cell types and tissues
CC and which may be involved in the maturation of CD34+ cells; certain
CC concentrations of SZF1-2 can cause hematopoietic cells to
CC differentiate and mature rather than to merely replicate. The
CC identification of SZF1 was made through the random sequencing of a
CC portion of several hundred clones from a cDNA library prepared from
CC human bone marrow CD34+ cells. One of the fragments was unique but
CC showed homology to the Kruppel family of zinc finger proteins.
CC This fragment was used to screen cDNA libraries from CD34+ cells,
CC K562 cells and human lung. Overlapping cDNA fragments which
CC hybridised to the probe gave rise to 2 alternatively spliced cDNA
CC products, termed SZF1-1 (see AA04360) and SZF1-2. The gene is
CC found on chromosome 3. Database searches showed the most highly
CC related genes to be ZNF13, KId1 and ZNF95 (65%, 55% and 45%
CC homology at the nucleotide level, respectively). The invention
CC additionally provides recombinant vectors and host cells, and a
CC method of producing SZF1 polypeptides. The polynucleotides may be
CC utilised for gene therapy in a host to replace or supplement a
CC defective SZF1 gene.

CC
CC
CC Sequence 3099 BP: 836 A; 744 C; 812 G; 707 T; 0 other;

SQ

XX

Query Match 1.0%; Score 27; DB 20; Length 3099;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2149 tcatcagaccacgagagacacacacag 2175
|||||
Db 1913 tcatcagaccacgagagacacacacag 1939

RESULT 17
ABL00474/c
ID ABL00474 standard; DNA; 51 BP.
XX
AC ABL00474;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:465.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytosstatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
OS
PN WO200138586-A2.
PN
PD 31-MAY-2001.
PD
PF 22-NOV-2000; 2000WO-US32311.
PF
PR 24-NOV-1999; 99US-0167383.
PR
PA (CURA-) CURAGEN CORP.
PA
PI Shimkets RA, Leach M;
PI WPI; 2001-355949/37.
XX
XX Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
XX Claim 1; Page 388; 674pp; English.

CC	ABLO00010 to ABL01104 represent human nucleic acid oligonucleotides comprising one or more single nucleotide polymorphisms (SNPs). ABB56531 to ABB56903 represent human peptides encoded by some of the SNP oligonucleotides. The sequences from the present invention can have immunosuppressive, cytostatic, antiinflammatory, neuroprotective and antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides and antibodies from the present invention can be used for treating a subject suffering from, at risk for, or suspected of, suffering from a pathology ascribed to the presence of a sequence polymorphism. The pathology may be autoimmune diseases, inflammation, cancer, diseases of the nervous system, and infection by pathogenic microorganisms. The SNPs are also useful for determining which forms of a characterised polymorphism are present in individuals. The antibodies may be used in the detection, quantitation and/or cellular or tissue localisation of a polymorphic protein (e.g., for use in measuring levels of the polymorphic protein within appropriate physiological samples).
CC	Sequence 51 BP; 9 A; 10 C; 14 G; 18 T; 0 other;
SQ	
Query Match	0.9%; Score 25; DB 23; Length 51; Best Local Similarity 100.0%; Pred. No. 0.076;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1378 gcttagccagaagtcacacctcat 1402 DB 51 GCCTTAGGCCAGACGTCACACTCAT 27
RESULT 18	
ABL00878/C	
ID	ABL00878 standard; DNA; 51 BP.
XX AC	ABL00878;
XX DT	05-MAR-2002 (first entry)
XX DE	Human amino acid change SNP oligonucleotide SEQ ID NO:869.
XX KM	Human; single nucleotide polymorphism; SNP; polymorphism; cytosstatic; immunosuppressive; antinflammatory; neuroprotective; antimicrobial; autoimmune disease; inflammation; cancer; nervous system disease; infection; polymorphic protein; ds.
KW XX	Homo sapiens.
OS XX	WO200138586-AZ.
PN XX	
PD XX	31-MAY-2001.
PX XX	22-NOV-2000; 2000WO-US32311.
PR XX	24-NOV-1999; 99US-0167383.
PA XX	(CURA-) CURAGEN CORP.
PI XX	Shinkets RA, Leach M;
PT XX	WP1, 2001-355949/37.
DR XX	
XX PT	Isolated human nucleic acids comprising one or more single nucleotide polymorphisms, useful for treating a subject suffering from a pathology, e.g. autoimmune diseases, ascribed to the presence of a sequence polymorphism -
PT XX	
PS XX	Claim 1; Page 509; 674pp; English.
XX CC	ABI00010 to ABL01104 represent human nucleic acid oligonucleotides comprising one or more single nucleotide polymorphisms (SNPs). ABB56531 to ABB56903 represent human peptides encoded by some of the SNP oligonucleotides. The sequences from the present invention can have immunosuppressive, cytostatic, antiinflammatory, neuroprotective and

CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).

XX
SQ Sequence 51 BP; 10 A; 14 C; 12 G; 15 T; 0 other;

Query Match 0.9%; Score 25; DB 23; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.076; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

OY 1362 agggagtgtgacgaagcttagcc 1386
|||||
Db 51 AGGAGTGTGACGACGCTTAGCC 27

RESULT 19
AAK92290
ID AAK92290 standard; cDNA; 763 BP.

XX AAK92290;

XX .06-NOV-2001 (first entry)

DE Human cDNA 5'-end sequence, SEQ ID NO: 750.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI: 2001-524255/58.

DR 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -

PS Claim 2; SEQ ID NO 750; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

SQ Sequence 763 BP; 169 A; 206 C; 217 G; 163 T; 8 other;

Query Match 0.9%; Score 25; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.074; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

OY 224 ggatgtgctgtgtactaccacag 248
|||||
Db 343 ggatgtgctgtgtactaccacag 367

RESULT 20
AAK93811
ID AAK93811 standard; cDNA; 763 BP.

XX AAK93811;

XX .06-NOV-2001 (first entry)

DE Human cDNA clone representative sequence, SEQ ID NO: 2271.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI: 2001-524255/58.

DR 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -

PS Example 11; SEQ ID NO 2271; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 763 BP; 169 A; 206 C; 217 G; 163 T; 8 other;

Query Match 0.9%; Score 25; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.074; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

OY 224 ggatgtgctgtgtactaccacag 248
|||||
Db 343 ggatgtgctgtgtactaccacag 367

RESULT 21
AAI95339
ID AAI95339 standard; CDNA; 814 BP.
XX
AC AAI95339;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 1414.
XX
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PN WO20016719-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01629.
XX
PR 07-MAR-2000; 2000JP-0159195.
XX
PA (CHIB-) CHIBA PREFECTURE.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
XX
PI Nakagawara A;
XX
DR WPI; 2001-565584/63.
XX
PT Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX
PS Claim 1; Page 1061; 2979pp; Japanese.
XX
CC The invention relates to novel genes (AAI93926-AAI97963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX
SQ Sequence 814 BP; 240 A; 170 C; 217 G; 170 T; 17 other;

Query Match 0.9%; Score 25; DB 22; Length 814;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1416 acacacacagagagagagccttatg 1440
DB 368 acacacacagagagagagccttatg 392
|||||

RESULT 22
ABA08823
ID ABA08823 standard; CDNA; 871 BP.
XX
AC ABA08823;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human PRO1847 homologue-encoding CDNA, SEQ ID NO:599.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;

KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritis; haemostatic; antiatherosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
XX
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
XX
DR P-PSDB; ABB11579.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 600; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides,
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 871 BP; 162 A; 258 C; 236 G; 215 T; 0 other;

Query Match 0.9%; Score 25; DB 22; Length 871;

Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 ggatggtgctgtgtactaccag 248
|||||
Db 336 ggatggtgctgtgtactaccag 360

RESULT 23

AAK94575
ID AAK94575 standard; cDNA; 2200 BP.

XX AAK94575;

XX 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 3494.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

DR P-PSDB; AAK93641.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3494; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a full length

CC human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 2200 BP; 579 A; 558 C; 529 G; 534 T; 0 other;

Query Match 0.9%; Score 25; DB 22; Length 2200;

Best Local Similarity 100.0%; Pred. No. 0.074;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 ggatggtgctgtgtactaccag 248

|||||

Db 343 ggatggtgctgtgtactaccag 367

RESULT 24

AAK91560

ID AAK91560 standard; cDNA; 688 BP.

XX AAK91560;

XX 06-NOV-2001 (first entry)

DE Human cDNA 5'-end sequence, SEQ ID NO: 20.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 2; SEQ ID NO 20; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is the nucleotide

CC sequence of the 5'-end of a cDNA provided in the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 688 BP; 156 A; 146 C; 173 G; 210 T; 3 other;

Query Match 0.8%; Score 23; DB 22; Length 688;

Best Local Similarity 100.0%; Pred. No. 0.72;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 214 tggcatcaggagatggtctg 236

|||||

Db 361 tggcatcaggagatggtctg 383

RESULT 25

AAK93129

ID AAK93129 standard; cDNA; 688 BP.

XX AAK93129;

XX 06-NOV-2001 (first entry)

DE Human cDNA clone representative sequence, SEQ ID NO: 1589.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

XX EPI130094-A2.

```
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX 11-JAN-2000; 2000JP-0118774.
PR
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation.
PS
XX Example 11; SEQ ID NO 1589; 1380bp + sequence listing; English.
PS
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used in
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 688 BP; 156 A; 146 C; 173 G; 210 T; 3 other;
SQ
Query Match 0.8%; Score 23; DB 22; Length 688;
Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 214 tggcattcaggagatgtgctgtc 236
DB 361 tggcattcaggagatgtgctgtc 383
RESULT 26
ABA50438/C
ID ABA50438 standard; DNA; 1512 BP.
XX
XX ABA50438;
AC
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Human breast cell single exon nucleic acid probe #9133.
DE
XX Human: microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00662.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
```

```
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-496933/54.
DR
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 9133; 327bp + sequence listing; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
SQ
Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1490 acaccagaggacacactcaggag 1512
DB 983 ACACCAGAGGACACACTCAGGG 961
RESULT 27
ABA68388/C
ID ABA68388 standard; DNA; 1512 BP.
XX
XX ABA68388;
AC
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Human foetal liver single exon nucleic acid probe #16693.
DE
XX Human: foetal liver; gene expression; single exon nucleic acid probe; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00669.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
```

XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human fetal liver -
 XX
 PS Claim 4; SEQ ID NO 16693; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
 Query Match 0.8%; Score 23; DB 22; Length 1512;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1490 acaccagagacactcagg 1512
 DB 983 ACACGAGGACACTCAGGG 961
 RESULT 28
 ID ABA35382/C
 XX ABA35382 standard; DNA; 1512 BP.
 XX
 AC ABA35382;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #13848 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX

PS Claim 4; SEQ ID No 13848; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
 Query Match 0.8%; Score 23; DB 22; Length 1512;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1490 acaccagagacactcagg 1512
 DB 983 ACACGAGGACACTCAGGG 961
 RESULT 29
 ID AAK16761/C
 XX AAK16761 standard; DNA; 1512 BP.
 XX
 AC AAK16761;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 16752.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 16752; 650pp + sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the invention.
XX
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 30
AAK42537/c
ID AAK42537 standard; DNA; 1512 BP.

AC AAK42537;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 17094.

KW Human; bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 17094; 658bp + Sequence listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

CC the probes of the invention.
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 31
AAI23284/c
ID AAI23284 standard; DNA; 1512 BP.

AC AAI23284;

DT 12-OCT-2001 (first entry)

DE Probe #13217 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 13217; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes

CC (SENP). The present sequence is one such probe. The SENPs are derived

CC from human HeLa cells. The SENPs can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene

CC expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 32

AAI48604/c
ID AAI48604 standard; DNA; 1512 BP.

AC AAI48604;

DT 17-OCT-2001 (first entry)

DE Probe #17290 used to measure gene expression in human placenta sample.

```

XX KW Probe: microarray; human; placenta; antenatal diagnosis:
XX KM genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-48897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta -
XX PS Claim 25; SEQ ID No 17290; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match          0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaagg 1512
DB 983 ACACGAGAGCAGCAGCTCAGCGG 961

RESULT 33
AAI08925/c
ID AAI08925 standard; DNA; 1512 BP.
XX AC AAI08925;
XX DT 09-OCT-2001 (first entry)
XX DE Probe #8916 used to measure gene expression in human breast sample.
XX KW Probe: human; breast disease; breast cancer; development disorder; ss;
XX KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US00661.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.

```

```

PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-476286/51.
XX PT Novel single exon nucleic acid probe used to measuring gene expression
XX PT in a human breast -
XX PS Claim 25; SEQ ID No 8916; 322pp; English.
XX CC The present invention relates to novel single exon nucleic acid probes.
XX CC The present sequence is one such probe. The probes are useful for
XX CC measuring human gene expression in a human breast sample, where the probe
XX CC hybridises at high stringency to a nucleic acid expressed in the human
XX CC breast. The probes are useful for predicting, diagnosing, grading,
XX CC staging, monitoring and prognosing diseases of the human breast,
XX CC particularly those diseases with polygenic aetiology. The diseases
XX CC include: breast cancer, disorders of development, inflammatory diseases
XX CC of the breast, fibrocystic changes, proliferative breast disease and
XX CC non-carcinoma tumours.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match          0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaagg 1512
DB 983 ACACGAGAGCAGCAGCTCAGCGG 961

RESULT 34
AAK94635
ID AAK94635 standard; CDNA; 1623 BP.
XX AC AAK94635;
XX DT 06-NOV-2001 (first entry)
XX DE Human full-length CDNA, SEQ ID NO: 3609.
XX KW Human; full length CDNA; CDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-0114089.
XX PR 08-JUL-1999; 99JP-0194486.
XX PR 11-JAN-2000; 2000JP-0118774.
XX PR 02-MAY-2000; 2000JP-0183765.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Makematsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2001-524255/58.
XX

```

PT 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
XX
PS Claim 8; SEQ ID NO 3609; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 1623 BP; 389 A; 346 C; 376 G; 512 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1623;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0;

OY 214 tggcattcaggagatgtgctgtg 236
|||||
DB 361 tggcattcaggagatgtgctgtg 383

RESULT 35
ABA45305/C
ID ABA45305 standard; DNA; 1965 BP.
XX
AC ABA45305;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #4000.
XX
KW Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 1; SEQ ID NO 4000; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived

CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0;

OY 1490 acaccagagagacacactcaggag 1512
|||||
DB 1110 ACACCAGAGAGACACACTCAGGCG 1088

RESULT 36
ABA55794/C
ID ABA55794 standard; DNA; 1965 BP.
XX
AC ABA55794;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #4099.
XX
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX
XX Claim 1; SEQ ID NO 4099; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcagggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGGG 1088

RESULT 37
ABA25474/C
ID ABA25474 standard; DNA; 1965 BP.

XX ABA25474;

DT 23-JAN-2002 (first entry)

DE Probe #3940 for gene expression analysis in human heart cell sample.

XX Human: gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.

XX Homo sapiens.

PN WO200157274-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

DR Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

XX Claim 1; SEQ ID No 3940; 530pp; English.

PS The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for

CC predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the

CC probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcagggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGGG 1088

RESULT 38
AAK04015/C
ID AAK04015 standard; DNA; 1965 BP.

XX AAK04015;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 4006.

XX Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

XX Homo sapiens.

PN WO200157275-A2.

XX 30-JAN-2001; 2001WO-US00667.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

DR Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 4006; 650pp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcagggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGGG 1088

RESULT 39
AAK29501/C

ID AAK29501 standard; DNA; 1965 BP.
XX
AC AAK29501;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 4058.
XX
DE Human bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 4058; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaggag 1512
|||||
DB 1110 ACACGAGGACACACTCAGGGG 1088

RESULT 40
AA14072/C
ID AA14072 standard; DNA; 1965 BP.
XX
AC AA14072;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4005 for gene expression analysis in human cervical cell sample.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
OS Homo sapiens.
XX
PN WO200157278-A2.

XX
PD 09-AUG-2001.
XX
AC 30-JAN-2001; 2001WO-US00670.
XX
DT 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 4005; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaggag 1512
|||||
DB 1110 ACACGAGGACACACTCAGGGG 1088

RESULT 41
AA13452/C
ID AA13452 standard; DNA; 1965 BP.
XX
AC AA13452;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4138 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 4138; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaggg 1512
DB 1110 ACACCAAGGACACACTCAGGG 1088

RESULT 42
AAI03925/C
ID AAI03925 standard; DNA; 1965 BP.
XX
AC AAI03925;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #3916 used to measure gene expression in human breast sample.
XX
KW Probe: human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
PS Claim 25; SEQ ID No 3916; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.

CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcaggg 1512
DB 1110 ACACCAAGGACACACTCAGGG 1088

RESULT 43
AAS81041
ID AAS81041 standard; cDNA; 2634 BP.
XX
AC AAS81041;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16845.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PsDB; ABG16854.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 16845; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2634 BP; 747 A; 686 C; 713 G; 488 T; 0 other;

Query Match 0.8%; Score 23; DB 23; Length 2634;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0;

OY 214 tggcattcaggagtgtgctctg 236
DB 2 tggcattcaggagtgtgctctg 24

RESULT 44
AAF67397
ID AAF67397 standard; CDNA; 439 BP.

XX AAF67397;

DT 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 3159.

DE Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KM breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.

OS WO200102568-A2.

PN 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US18374.

XX 02-JUL-1999; 99US-0142310.

PR 02-JUL-1999; 99US-0142311.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirkeljakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LM, Strachan-Crain B;

XX WPI; 2001-091805/10.

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -

XX Claim 9; Page 1017; 1046pp; English.

CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in

CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX

SQ Sequence 439 BP; 104 A; 109 C; 130 G; 96 T; 0 other;

Query Match 0.8%; Score 22; DB 22; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

OY 228 gtgctgtgtacttaccaccag 249
DB 156 gtgctgtgtacttaccaccag 177

RESULT 45
AAC79234
ID AAC79234 standard; CDNA; 484 BP.

XX AAC79234;

DT 05-FEB-2001 (first entry)

XX Human lung tumour-specific CDNA #187.

XX Lung tumour protein; lung cancer; cytostatic; vaccine; ss.

XX Homo sapiens.

OS WO200060077-A2.

PN 12-OCT-2000.

XX 30-MAR-2000; 2000WO-US08560.

XX 02-APR-1999; 99US-0285323.

PR 09-AUG-1999; 99US-0370838.

PR 30-DEC-1999; 99US-0476235.

PR 03-MAR-2000; 2000US-0518809.

XX (CORI-) CORIXA CORP.

PA Reed SG, Lodes MJ, Mohamath R, Secrist H;

XX WPI; 2000-638466/61.

PT Novel lung tumor polypeptides and polynucleotides, useful for
PT detecting, monitoring or treating cancer, especially lung cancer -

XX Claim 3; Page 206; 243pp; English.

CC The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed.
CC The polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer.
XX

SQ Sequence 484 BP; 145 A; 106 C; 135 G; 98 T; 0 other;

Query Match 0.8%; Score 22; DB 21; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0;

OY 1089 ggggagaaaccttatgtgtgca 1110
|||||

Db 233 ggggagaacctatgtgtgca 254

RESULT 46
ID AAD23310 standard; cDNA: 484 BP.
XX
AC AAD23310;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific 2LT-120 cDNA.
XX
KW Human: lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 2LT-120; ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09991.
XX
PR 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
PA "(CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR WPI; 2001-639201/73.
XX
PT New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
XX
PS Example 4; Page 256-257; 378pp; English.
XX
CC The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is human lung tumour-specific cDNA.
XX
SO Sequence 484 BP; 145 A; 106 C; 135 G; 98 T; 0 other.

Query Match 0.8%; Score 22; DB 23; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1089 ggggagaacctatgtgtgca 1110
|||||
Db 233 ggggagaacctatgtgtgca 254

RESULT 47
ID AAS26521 standard; cDNA: 758 BP.
XX
AC AAS26521;
XX
DT 07-NOV-2001 (first entry)
XX

DE Human cDNA encoding a novel secreted protein, Seq ID 700.
XX
KW Human: immunosuppressive; antiarthritic; ss; antineumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.

[illegible]

PR	05-DEC-2000;	2000US-0251988.	
PR	05-DEC-2000;	2000US-0256719.	
PR	06-DEC-2000;	2000US-0251479.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251990.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
PA	(HUMA-)	HUMAN GENOME SCI INC.	
PI	Rosen CA,	Barash SC, Ruben SM;	
DR	WPI:	2001-488783/53.	
DR	P-PSDB:	AAU16534.	
XX			
PT	New nucleic acid molecules encoding 461 human secreted proteins for		
PT	diagnosing, preventing, treating or ameliorating medical conditions and		
PT	used as food additives or preservatives -		
XX			
PS	Claim 1;	SEQ ID NO 700; 980bp; English.	
XX			
CC	The invention relates to isolated nucleic acid molecules and their		
CC	encoded secreted proteins. The nucleic acids and proteins are used to		
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They		
CC	are also used in diagnosing a pathological condition or susceptibility		
CC	to a pathological condition. Antibodies to the proteins can also		
CC	be used in alleviating symptoms associated with the disorders and in		
CC	diagnostic immunoassays e.g. radioimmunoassays or enzyme linked		
CC	immunosorbent assays (ELISA). Disorders which are diagnosed or treated		
CC	include autoimmune diseases e.g. rheumatoid arthritis,		
CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,		
CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders		
CC	e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.		
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi		
CC	and ocular disorders e.g. corneal infection, and many other		
CC	disorders listed in the specification. The polypeptides can also		
CC	be used to aid wound healing and epithelial cell proliferation, to		
CC	prevent skin aging due to sunburn, to maintain organs before		
CC	transplantation, for supporting cell culture of primary tissues, to		
CC	regenerate tissues and in chemotaxis. The polypeptides can also be used		
CC	as a food additive or preservative to increase or decrease storage		
CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,		
CC	minerals, cofactors and other nutritional components. The present		
CC	sequence encodes a novel secreted protein of the invention.		
Query Match	0.8%;	Score 22;	DB 22; Length 758;
Best Local Similarity	100.0%;	Pred. No. 2.2;	
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1419	cacacagagagaagccttatg	1440
DB	283	cacacagagagaagccttatg	304
RESULT 48			
ABA21456			
ID	ABA21456 standard; DNA; 1038 BP.		
XX			
AC	ABA21456;		
XX			
DT	23-JAN-2002 (first entry)		
XX			
DE	Human nervous system related polynucleotide seq ID NO 13787.		
XX			
KW	Human; nootropic; neuroprotective; cytosatic; dermatological; virocidic;		
KW	immunosuppressive; antinflammatory; anti-HIV; antibacterial; antiviral;		
KW	antiparkinsonian; antisticking; antianemic; antiarthritic; cancer;		
KW	antirheumatic; hepatotropic; cerebroprotective; antinflammatory;		
KW	antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;		

KW anti-parasitic; cardiac; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001MO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0203515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235534.
 PR 27-SEP-2000; 2000US-0235536.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249266.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.

CC	XX	PR	20-OCT-2000; 2000US-0241221.
CC	XX	PR	20-OCT-2000; 2000US-0241785.
CC	XX	PR	20-OCT-2000; 2000US-0241786.
CC	XX	PR	20-OCT-2000; 2000US-0241787.
CC	XX	PR	20-OCT-2000; 2000US-0241808.
CC	XX	PR	20-OCT-2000; 2000US-0241809.
CC	XX	PR	20-OCT-2000; 2000US-0241826.
CC	XX	PR	01-NOV-2000; 2000US-0244617.
CC	XX	PR	08-NOV-2000; 2000US-0246474.
CC	XX	PR	08-NOV-2000; 2000US-0246475.
CC	XX	PR	08-NOV-2000; 2000US-0246476.
CC	XX	PR	08-NOV-2000; 2000US-0246477.
CC	XX	PR	08-NOV-2000; 2000US-0246478.
CC	XX	PR	08-NOV-2000; 2000US-0246523.
CC	XX	PR	08-NOV-2000; 2000US-0246524.
CC	XX	PR	08-NOV-2000; 2000US-0246525.
CC	XX	PR	08-NOV-2000; 2000US-0246526.
CC	XX	PR	08-NOV-2000; 2000US-0246527.
CC	XX	PR	08-NOV-2000; 2000US-0246528.
CC	XX	PR	08-NOV-2000; 2000US-0246532.
CC	XX	PR	08-NOV-2000; 2000US-0246609.
CC	XX	PR	08-NOV-2000; 2000US-0246610.
CC	XX	PR	08-NOV-2000; 2000US-0246611.
CC	XX	PR	08-NOV-2000; 2000US-0246613.
CC	XX	PR	17-NOV-2000; 2000US-0249207.
CC	XX	PR	17-NOV-2000; 2000US-0249208.
CC	XX	PR	17-NOV-2000; 2000US-0249209.
CC	XX	PR	17-NOV-2000; 2000US-0249210.
CC	XX	PR	17-NOV-2000; 2000US-0249211.
CC	XX	PR	17-NOV-2000; 2000US-0249212.
CC	XX	PR	17-NOV-2000; 2000US-0249213.
CC	XX	PR	17-NOV-2000; 2000US-0249214.
CC	XX	PR	17-NOV-2000; 2000US-0249215.
CC	XX	PR	17-NOV-2000; 2000US-0249216.
CC	XX	PR	17-NOV-2000; 2000US-0249217.
CC	XX	PR	17-NOV-2000; 2000US-0249218.
CC	XX	PR	17-NOV-2000; 2000US-0249244.
CC	XX	PR	17-NOV-2000; 2000US-0249245.
CC	XX	PR	17-NOV-2000; 2000US-0249264.
CC	XX	PR	17-NOV-2000; 2000US-0249265.
CC	XX	PR	17-NOV-2000; 2000US-0249297.
CC	XX	PR	17-NOV-2000; 2000US-0249300.
CC	XX	PR	17-NOV-2000; 2000US-0249309.
CC	XX	PR	01-DEC-2000; 2000US-0250160.
CC	XX	PR	01-DEC-2000; 2000US-0250391.
CC	XX	PR	05-DEC-2000; 2000US-0251030.
CC	XX	PR	05-DEC-2000; 2000US-0251988.
CC	XX	PR	05-DEC-2000; 2000US-0256719.
CC	XX	PR	06-DEC-2000; 2000US-0251479.
CC	XX	PR	08-DEC-2000; 2000US-0251856.
CC	XX	PR	08-DEC-2000; 2000US-0251858.
CC	XX	PR	08-DEC-2000; 2000US-0251869.
CC	XX	PR	08-DEC-2000; 2000US-0251989.
CC	XX	PR	08-DEC-2000; 2000US-0251990.
CC	XX	PR	11-DEC-2000; 2000US-0254097.
CC	XX	PR	05-JAN-2001; 2001US-0259678.
CC	XX	PA	(HUMA-) HUMAN GENOME SCI INC.
CC	XX	PI	Rosen CA, Barash SC, Ruben SH;
CC	XX	DR	WPI; 2001-483426/52.
CC	XX	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
CC	XX	PT	useful for preventing, diagnosing and/or treating cancers and
CC	XX	PT	metastasis -
CC	XX	PS	Disclosure; SEQ ID NO 39080; 307bp + Sequence Listing; English.
CC	XX	XX	AANK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC	XX	XX	antigen acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
CC	XX	XX	activity, and can be used in gene therapy and vaccine production. (II)
CC	XX	XX	proteins and polynucleotides may be used in the prevention, diagnosis and

CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	supplement the patients own production of (I).
CC	Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I), by inserting
CC	the nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/hematopoietic-related diseases, especially
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC	to AAK87694 represent human immune/hematopoietic antigen genomic
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM62169
CC	represent sequences used in the exemplification of the present invention.
XX	
SO	Sequence 1038 BP; 238 A; 242 C; 184 G; 374 T; 0 other;
Query Match	0.8%; Score 22; DB 22; Length 1038;
Best Local Similarity	100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Gy	1251 cacactggggaagacattatg 1272
Db	843 CACACTGGGAGAGCCTTATG 822
RESULT 50	
AA571465	
ID	AA571465 standard; cDNA; 1185 BP.
XX	
AC	AA571465;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #7269.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	Food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
XX	
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEO INC.
XX	
P1	Drmnac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
XX	
PT	P-PSDB: ABG07278.
XX	
PS	Claim 1; SEQ ID NO 7269; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

SQ Sequence 1185 BP; 344 A; 262 C; 295 G; 284 T; 0 other;

Query Match 0.8%; Score 22; DB 23; Length 1185;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1587 cacacgagggaagaccattg 1608
 |||
 DB 139 cacacgagggaagaccattg 160

Search completed: May 16, 2002, 10:45:05
 Job time: 8490 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:20:35 ; Search time 2877.72 Seconds
(without alignments)
13001.112 Million cell updates/sec

Title: US-09-898-556A-3

Perfect score: 2772

Sequence: 1 cagcgcgcgttaagctggtt.....ttctaccaccctcaccct 2772

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 674847542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	721	26.0	1082	10	BM450709 AGENCOURT
2	677	24.4	1035	10	BM470740 AGENCOURT
3	585	21.1	1071	10	BE741389
4	581	21.0	590	9	AW245709
5	545	19.7	1009	10	BM449478
6	510	18.4	563	9	AI669533
7	509	18.4	535	9	AI047439
8	492	17.7	537	9	AI802142
9	482	17.4	537	9	BE220142
10	471	17.0	624	9	BE178716
11	458	16.5	512	10	BE111754
12	454	16.4	505	10	BE463501
13	449	16.2	486	9	AW401692
14	419	15.1	486	9	AW720986
15	414	14.9	613	9	AW364787
16	413	14.9	697	10	BE391140
17	405	14.6	673	10	BC928720

C 18	402	14.5	453	9	AW513018	AW513018 xct6f03.x
C 19	399	14.4	481	9	AI276016	AI276016 qw08h02.x
C 20	395	14.2	451	9	AA450184	AA450184 2x42e09.x
C 21	383	13.8	435	9	AI081665	AI081665 ou63e06.s
C 22	383	13.8	449	9	A1750087	A1750087 alc35f07.x
C 23	378	13.6	455	10	BE501382	BE501382 hws1a02.x
C 24	378	13.6	728	10	BE746403	BE746403 601579525
C 25	372	13.4	424	9	AI016684	AI016684 ou96d05.x
C 26	372	13.4	458	9	AI440109	AI440109 t156h06.x
C 27	372	13.4	428	9	AI953623	AI953623 wg23a03.x
C 28	366	13.2	501	9	AW140124	AW140124 UI-H-BIL-
C 29	363	13.1	426	9	AI828662	AI828662 tr06f04.x
C 30	355	12.8	486	9	AI650983	AI650983 wa96d07.x
C 31	351	12.7	473	10	BE476173	BE476173 naa29c11.
C 32	350	12.6	385	9	AA218658	AA218658 zq96g04.r
C 33	350	12.6	500	9	AI986444	AI986444 wz65d09.x
C 34	345	12.4	551	10	BE395838	BE395838 601310077
C 35	344	12.4	508	9	AI860707	AI860707 w115d12.x
C 36	343	12.4	399	9	AI289399	AI289399 qw32e05.x
C 37	340	12.3	392	9	AA218659	AA218659 zq96g04.s
C 38	340	12.3	395	9	AI361705	AI361705 qz18f04.x
C 39	337	12.2	513	9	AW958936	AW958936 EST371006
C 40	333	12.0	794	10	BI256434	BI256434 602974454
C 41	332	12.0	383	9	AA450118	AA450118 2x42e09.s
C 42	318	11.5	378	9	AA484745	AA484745 ne81h06.s
C 43	315	11.4	369	9	AI159662	AI159662 qb80f10.x
C 44	315	11.4	818	9	AL566746	AL566746 AL566746
C 45	312	11.3	501	10	R50617	R50617 yj57b01.r1
C 46	311	11.2	427	10	R83624	R83624 yp16h01.r1
C 47	308	11.1	410	9	AW513538	AW513538 x086b03.x
C 48	305	11.0	383	9	AW139497	AW139497 UI-H-BIL-
C 49	291	10.5	680	9	AW579734	AW579734 RC1-HT025
C 50	290	10.5	1041	10	BE798936	BE798936 601583719
C 51	279	10.1	712	10	BF026155	BF026155 601668851
C 52	279	10.1	761	11	BC010024	BC010024 Home sapi
C 53	269	9.7	809	10	BF746802	BF746802 602704138
C 54	256	9.2	472	10	BF844319	BF844319 RC1-HT022
C 55	255	9.2	275	9	AW360766	AW360766 PM3-CT024
C 56	251	9.1	577	10	BE275642	BE275642 601121171
C 57	249	9.0	884	10	BG402277	BG402277 602465851
C 58	246	8.9	323	9	AI655251	AI655251 wd68a12.x
C 59	246	8.9	474	10	BF940868	BF940868 hu60b12.x
C 60	240	8.7	377	10	BF773364	BF773364 CM2-IT003
C 61	239	8.6	669	10	BE618307	BE618307 601462687
C 62	228	8.2	439	9	AW753060	AW753060 PM3-CT024
C 63	223	8.0	481	9	AW574555	AW574555 UI-HR-BKO
C 64	223	8.0	565	9	AA777271	AA777271 z196b07.s
C 65	215	7.8	451	10	BF768400	BF768400 QV3-IT000
C 66	209	7.5	317	9	AI917803	AI917803 wd53b08.x
C 67	209	7.5	456	9	AI393883	AI393883 t965h06.x
C 68	207	7.5	260	9	AI420536	AI420536 te99a06.x
C 69	205	7.4	530	9	AW954105	AW954105 EST366175
C 70	204	7.3	371	9	AI633088	AI633088 t201h05.x
C 71	202	7.3	371	9	AW868888	AW868888 MK1-SN006
C 72	198	7.1	300	9	AA384543	AA384543 EST98201
C 73	198	7.1	524	10	R90992	R90992 yp93b08.r1
C 74	198	7.1	267	9	AL537287	AL537287 AL537287
C 75	197	7.1	908	9	AI990117	AI990117 ws29d01.x
C 76	195	7.0	276	9	AW134770	AW134770 UI-H-BIL-
C 77	189	7.0	281	9	AA012980	AA012980 ze35h03.s
C 78	189	6.8	664	10	BE619323	BE619323 601473319
C 79	182	6.6	304	9	AI868995	AI868995 wd91h09.x
C 80	179	6.5	213	10	BC995316	BC995316 CM0-HT129
C 81	171	6.2	180	10	BI004708	BI004708 CM0-HN020
C 82	160	5.8	324	10	BF222885	BF222885 7q27a04.x
C 83	157	5.7	450	9	BE167882	BE167882 CM1-HT051
C 84	156	5.6	401	10	BC010675	BC010675 CM4-GN029
C 85	155	5.6	252	9	AI937720	AI937720 wp83b12.x
C 86	155	5.6	319	9	AW051156	AW051156 w208d02.x
C 87	152	5.5	326	10	BE352587	BE352587 IL3-HT061
C 88	149	5.4	200	10	BC995318	BC995318 CM0-HT129
C 89	147	5.3	260	10	BI006722	BI006722 CM3-RT000
C 90	146	5.3	599	10	BC470016	BC470016 602533494

529	20	0.7	1823	11	AK006349	AK006349 Mus muscu	602	19	0.7	418	9	AM058241	AM058241 wx17f03.x
530	20	0.7	2446	11	AK014064	AK014064 Mus muscu	603	19	0.7	418	10	BF601786	BF601786 266848 MA
531	20	0.7	2737	11	AK020267	AK020267 Homo sapi	604	19	0.7	423	10	AV722193	AV722193 av722193
532	20	0.7	2785	11	BC013395	BC013395 Homo sapi	605	19	0.7	423	10	R59875	R59875 yno7d410.r1
533	19	0.7	133	10	BE449468	BE449468 EST356227	606	19	0.7	424	9	A1274165	A1274165 q157g12.x
534	19	0.7	154	10	BE709019	BE709019 QV2-Ht057	607	19	0.7	427	10	R49560	R49560 y668a10.s1
535	19	0.7	156	10	BE709019	BE709019 QV2-Ht057	608	19	0.7	430	9	AF150148	AF150148 AF150148
536	19	0.7	192	10	BM483654	BM483654 55564 MAR	609	19	0.7	432	10	W05057	W05057 za78b10.r1
537	19	0.7	193	10	AM481765	AM481765 RC3-CN001	610	19	0.7	434	9	BM242363	BM242363 BM242363
538	19	0.7	201	10	BM316907	BM316907 947027H08	611	19	0.7	434	10	BM366743	BM366743 BM366743
539	19	0.7	227	10	NA0120	NA0120 yw72e12.r1	612	19	0.7	436	9	AM341477	AM341477 hd10b12.x
540	19	0.7	234	9	AV314795	AV314795 AV314795	613	19	0.7	437	10	HI7015	HI7015 ym39b12.r1
541	19	0.7	244	10	BI030393	BI030393 IL5-MT026	614	19	0.7	440	9	AA546021	AA546021 vk61a07.r
542	19	0.7	248	9	BB036384	BB036384 BB036384	615	19	0.7	441	9	AA043359	AA043359 zk62e11.r
543	19	0.7	249	9	AL644722	AL644722 AL644722	616	19	0.7	441	9	A1300433	A1300433 q013e03.x
544	19	0.7	251	10	WB4990	WB4990 mF43g11.r1	617	19	0.7	442	9	AA760366	AA760366 vv76b03.r
545	19	0.7	261	10	BC014593	BC014593 PM4-GN030	618	19	0.7	442	10	BF720277	BF720277 mab51e01.
546	19	0.7	262	10	BI040765	BI040765 CM3-NT026	619	19	0.7	443	9	A1528907	A1528907 ud42g08.y
547	19	0.7	279	9	A1685676	A1685676 tL89g07.x	620	19	0.7	446	12	AA0495847	AA0495847 HS_5212_B
548	19	0.7	284	12	AZ318450	AZ318450 IM0037L24	621	19	0.7	447	9	A1373029	A1373029 q202h05.x
549	19	0.7	287	9	BB232707	BB232707 BB232707	622	19	0.7	449	9	A1185656	A1185656 qe65c08.x
550	19	0.7	288	10	Z44630	Z44630 HSC25H081.n	623	19	0.7	449	10	BF080636	BF080636 231519 MA
551	19	0.7	295	12	A2862852	A2862852 2M0170A15	624	19	0.7	450	10	BF463158	BF463158 UT-R-CC0P
552	19	0.7	300	9	AU100566	AU100566 AU100566	625	19	0.7	453	9	BB838992	BB838992 BB838992
553	19	0.7	302	9	AM315867	AM315867 13583 MAR	626	19	0.7	453	10	BF719163	BF719163 mab37b06.
554	19	0.7	305	9	BE111253	BE111253 UT-R-BT1-	627	19	0.7	455	9	BB819591	BB819591 BB819591
555	19	0.7	306	9	AM176121	AM176121 IL0-BT016	628	19	0.7	459	10	HI6701	HI6701 ym26f10.s1
556	19	0.7	306	9	BB140184	BB140184 BB140184	629	19	0.7	459	10	BE419737	BE419737 UT-R-CM0-
557	19	0.7	307	9	BB395690	BB395690 BB395690	630	19	0.7	460	9	AL038051	AL038051 DFE2P56H
558	19	0.7	309	9	BB807689	BB807689 BB807689	631	19	0.7	462	10	BM365924	BM365924 BS3200200
559	19	0.7	310	10	BI307127	BI307127 PTSS0450	632	19	0.7	462	12	AA052586	AA052586 RPCI11-52
560	19	0.7	313	10	BI306987	BI306987 PTSS0281	633	19	0.7	463	9	AA043360	AA043360 zk62e11.s3
561	19	0.7	326	10	FI0244	FI0244 HSC3CD012.n	634	19	0.7	463	9	AA468863	AA468863 hd29b08.x
562	19	0.7	332	9	AA560291	AA560291 v167m05.r	635	19	0.7	463	10	BF386449	BF386449 UT-R-CA1-
563	19	0.7	334	10	W07223	W07223 za94d07.r1	636	19	0.7	463	10	BM147542	BM147542 TCAP1012
564	19	0.7	337	12	AQ239297	AQ239297 RPCI11-67	637	19	0.7	465	12	AQ340817	AQ340817 HS_2251_B
565	19	0.7	341	10	BF149452	BF149452 uy65b12.y	638	19	0.7	467	10	BF815230	BF815230 RC5-C101A
566	19	0.7	347	10	T82354	T82354 yc14b12.r1	639	19	0.7	468	9	AV742262	AV742262 AV742262
567	19	0.7	353	12	AQ376637	AQ376637 RPCI-11-1	640	19	0.7	468	10	BM106448	BM106448 510183 MA
568	19	0.7	357	9	AA562950	AA562950 v157b11.r	641	19	0.7	470	10	BM220898	BM220898 C0947R01-
569	19	0.7	360	9	AA740344	AA740344 OB23802.S	642	19	0.7	470	12	AQ806107	AQ806107 HS_3207_B
570	19	0.7	360	9	AL546726	AL546726 AL546726	643	19	0.7	472	9	AA138214	AA138214 mq76f07.r
571	19	0.7	365	10	BE773795	BE773795 QV1-F7016	644	19	0.7	472	12	AQ927284	AQ927284 RPCI-23-2
572	19	0.7	367	12	AZ499158	AZ499158 IM0136H19	645	19	0.7	473	9	AA853994	AA853994 a51h12.s
573	19	0.7	369	9	AA971411	AA971411 OP2C10.S	646	19	0.7	474	9	A1892950	A1892950 mq76f07.y
574	19	0.7	369	10	BB836473	BB836473 QV0-FN006	647	19	0.7	474	10	W85779	W85779 zh55h11.r1
575	19	0.7	370	9	AM527624	AM527624 UT-R-BT1-	648	19	0.7	476	9	AM408135	AM408135 UT-R-BM0
576	19	0.7	378	9	AM481123	AM481123 34467 MAR	649	19	0.7	478	10	N80080	N80080 yz87b05.r1
577	19	0.7	379	10	BI293236	BI293236 UT-R-DK0-	650	19	0.7	480	9	AA130914	AA130914 z014a04.r
578	19	0.7	379	10	BE773610	BE773610 QV1-F7016	651	19	0.7	482	9	A1291283	A1291283 gm16f09.x
579	19	0.7	379	12	AQ425420	AQ425420 C1PH1-EI-	652	19	0.7	482	10	W12166	W12166 m552h09.r1
580	19	0.7	382	9	A1088612	A1088612 QB14607.x	653	19	0.7	483	10	T81415	T81415 yd96c09.s1
581	19	0.7	382	12	AQ167216	AQ167216 HS_5171_A	654	19	0.7	483	12	AZ220047	AZ220047 IM0039120
582	19	0.7	385	9	AA184558	AA184558 mF50c05.r	655	19	0.7	485	10	BM288561	BM288561 529868 MA
583	19	0.7	385	10	BF765270	BF765270 IL2-CS005	656	19	0.7	487	10	BF002780	BF002780 7960h09.x
584	19	0.7	392	9	AM576046	AM576046 UT-HR-BL0	657	19	0.7	490	12	AZ102762	AZ102762 RPCI-23-4
585	19	0.7	396	9	A1239643	A1239643 qH32C06.x	658	19	0.7	491	10	R53168	R53168 y86bh03.r1
586	19	0.7	397	10	BC992384	BC992384 MR3-HT110	659	19	0.7	492	10	BT550419	BT550419 603192482
587	19	0.7	398	12	AQ790712	AQ790712 HS_3136_B	660	19	0.7	493	9	AM669714	AM669714 113237 MA
588	19	0.7	398	12	AZ307981	AZ307981 IM0010C06	661	19	0.7	493	12	AZ702835	AZ702835 RPCI-23-1
589	19	0.7	398	12	AQ194134	AQ194134 RPCI11-61	662	19	0.7	495	9	A1980622	A1980622 pat-PK002
590	19	0.7	398	12	AQ020576	AQ020576 RPCI11-58	663	19	0.7	496	9	A1740708	A1740708 wg17c07.x
591	19	0.7	399	9	A1641809	A1641809 v157b11.y	664	19	0.7	497	9	A1376339	A1376339 EST000061
592	19	0.7	405	9	A1939425	A1939425 q202g06.x	665	19	0.7	497	10	BC494561	BC494561 602540605
593	19	0.7	408	12	AQ323781	AQ323781 RPCI11-10	666	19	0.7	497	10	BF420611	BF420611 UT-R-B02
594	19	0.7	409	9	A1348516	A1348516 q034f03.x	667	19	0.7	498	10	BM364844	BM364844 BS3200060
595	19	0.7	409	12	AZ500537	AZ500537 o4e06fs.r	668	19	0.7	499	9	AM613309	AM613309 hg87e09.x
596	19	0.7	411	10	BI201093	BI201093 o4e06fs.r	669	19	0.7	500	10	BC613391	BC613391 def38e01.
597	19	0.7	412	12	CNS03BD2	AL554495 Tetradon	670	19	0.7	502	12	AQ322866	AQ322866 RPCI11-10
598	19	0.7	413	12	AF049571	AF049571 AF049571	671	19	0.7	505	9	AM534184	AM534184 UT-R-C4-a
599	19	0.7	415	12	AQ323224	AQ323224 RPCI11-10	672	19	0.7	506	10	BE476033	BE476033 158584 BA
600	19	0.7	415	12	AQ518301	AQ518301 HS_5105_A	673	19	0.7	506	12	AZ234941	AZ234941 RPCI-23-9
601	19	0.7	416	9	AM012164	AM012164 um08f01.y	674	19	0.7	506	12	AO594268	AO594268 HS_2095_A

675	19	0.7	511	10	BG69920	BG69920 602681284	c 748	19	0.7	602	9	A1811577	A1811577 tw74905.x
676	19	0.7	511	10	BE30939	BE30939 601094515	c 749	19	0.7	604	9	AM362174	AM362174 PMO-CT026
677	19	0.7	511	12	AO203360	AO203360 HS_2180.B	c 750	19	0.7	606	10	BP955407	BP955407 MR4-NN119
678	19	0.7	512	9	AM500026	AM500026 UT-HF-BNO	c 751	19	0.7	609	10	BE613897	BE613897 601504016
679	19	0.7	515	9	A1351041	A1351041 qt22a06.x	c 752	19	0.7	610	9	AU129289	AU129289 A1229289
680	19	0.7	516	10	BE665069	BE665069 153029.MA	c 753	19	0.7	612	9	AW390354	AW390354 RC0-ST018
681	19	0.7	520	12	AO442373	AO442373 HS_5117.B	c 754	19	0.7	617	10	BJ063466	BJ063466 BJ063466
682	19	0.7	521	9	AV707813	AV707813 AV707813	c 755	19	0.7	620	10	W28026	W28026 40H9 Human
683	19	0.7	522	10	BF894019	BF894019 OV1-MT016	c 756	19	0.7	628	12	A2365678	A2365678 A12365678
684	19	0.7	524	10	A1552538	A1552538 vx23f02.x	c 757	19	0.7	628	12	BE084068	BE084068 PMO-BT065
685	19	0.7	524	12	AO320430	AO320430 RPTC11-08	c 758	19	0.7	629	12	A2369192	A2369192 1M0119H10
686	19	0.7	526	12	AO133104	AO133104 HS_2188.A	c 759	19	0.7	632	10	BF453362	BF453362 mae67h12.
687	19	0.7	527	10	BF845977	BF845977 QV1-EN005	c 760	19	0.7	633	9	BB646933	BB646933 BB646933
688	19	0.7	527	10	BF015717	BF015717 uy28B02.Y	c 761	19	0.7	634	12	BB659863	BB659863 BB659863
689	19	0.7	527	10	BF332074	BF332074 PMO-BT065	c 762	19	0.7	634	12	A2422106	A2422106 1M0200H17
690	19	0.7	528	10	BG336018	BG336018 naf21b07.	c 763	19	0.7	634	12	BH293106	BH293106 CT230-44E
691	19	0.7	532	12	AO719865	AO719865 HS_5546.A	c 764	19	0.7	635	12	A2318784	A2318784 1M0037A23
692	19	0.7	533	9	AV668145	AV668145 AV668145	c 765	19	0.7	636	9	BB645228	BB645228 BB645228
693	19	0.7	534	12	A2329752	A2329752 1M0054F15	c 766	19	0.7	636	10	BG256784	BG256784 602371007
694	19	0.7	534	12	A2515060	A2515060 1M0054F15	c 767	19	0.7	637	12	AO590728	AO590728 HS_2094.B
695	19	0.7	535	9	AV609197	AV609197 AV609197	c 768	19	0.7	638	9	BB624175	BB624175 BB624175
696	19	0.7	536	9	AV609198	AV609198 AV609198	c 769	19	0.7	638	10	BM487139	BM487139 PGM2h-PK0
697	19	0.7	536	10	BF906287	BF906287 IL3-MT026	c 770	19	0.7	639	10	BE269572	BE269572 601184918
698	19	0.7	538	9	AV603850	AV603850 AV603850	c 771	19	0.7	639	10	BE294572	BE294572 601176145
699	19	0.7	539	9	BF435065	BF435065 7P05c07.x	c 772	19	0.7	650	10	BE533617	BE533617 601235242
700	19	0.7	541	12	AO504340	AO504340 RPTC11-2	c 773	19	0.7	650	12	A2939633	A2939633 2M0198M11
701	19	0.7	543	9	AA123106	AA123106 na24e05.F	c 774	19	0.7	652	9	AM985760	AM985760 ut85e12.Y
702	19	0.7	543	9	AM820820	AM820820 RC2-ST030	c 775	19	0.7	653	10	BG283656	BG283656 602407414
703	19	0.7	543	10	BC018287	BC018287 de66h11.Y	c 776	19	0.7	654	9	BB652891	BB652891 BB652891
704	19	0.7	544	9	AU149008	AU149008 AU149008	c 777	19	0.7	654	12	AO008409	AO008409 CIT-HSP-2
705	19	0.7	547	9	AV737616	AV737616 AV737616	c 778	19	0.7	656	12	AO627194	AO627194 CITBI-El-
706	19	0.7	547	10	BG362949	BG362949 sac14C07.	c 779	19	0.7	657	10	BE269529	BE269529 601184742
707	19	0.7	548	9	BE095874	BE095874 UT-R-BU0-	c 780	19	0.7	659	10	BG285524	BG285524 602380515
708	19	0.7	550	9	A1112304	A1112304 UT-R-Y0-m	c 781	19	0.7	659	10	BG754485	BG754485 602710112
709	19	0.7	550	10	BM128506	BM128506 1f12a02.Y	c 782	19	0.7	662	10	BF678357	BF678357 602084982
710	19	0.7	550	12	A2696034	A2696034 RPTC1-23-2	c 783	19	0.7	663	10	BE384954	BE384954 601276805
711	19	0.7	553	12	AO489671	AO489671 RPTC1-11-2	c 784	19	0.7	664	9	BB534363	BB534363 BB534363
712	19	0.7	556	12	A2877010	A2877010 2M0192G22	c 785	19	0.7	664	10	BE395137	BE395137 601310764
713	19	0.7	558	9	AM001430	AM001430 wu31d09.x	c 786	19	0.7	664	12	A2386848	A2386848 1M0146104
714	19	0.7	559	10	B1962420	B1962420 1d32g11.Y	c 787	19	0.7	665	10	BI151280	BI151280 602917250
715	19	0.7	559	10	BM287496	BM287496 5282g33.MA	c 788	19	0.7	666	10	BF000131	BF000131 7h18b10.x
716	19	0.7	561	10	BI714868	BI714868 1c33809.x	c 789	19	0.7	668	11	BC017760	BC017760 Homo sapi
717	19	0.7	563	9	AM295024	AM295024 UT-H-B12-	c 790	19	0.7	668	12	AG056194	AG056194 Pan trogl
718	19	0.7	569	9	AV595425	AV595425 AV595425	c 791	19	0.7	669	9	BB469375	BB469375 BB469375
719	19	0.7	571	12	FR0026225	FR0026225 F.rubripie	c 792	19	0.7	671	10	BF981509	BF981509 602309357
720	19	0.7	573	10	B1256649	B1256649 602973947	c 793	19	0.7	671	10	BI461982	BI461982 603203836
721	19	0.7	574	12	A2026858	A2026858 RPTC1-23-3	c 794	19	0.7	671	12	AC058350	AC058350 Pan trogl
722	19	0.7	574	10	BG390896	BG390896 602416956	c 795	19	0.7	672	12	AO998686	AO998686 RPTC1-23-3
723	19	0.7	575	10	BI962380	BI962380 1d32c06.Y	c 796	19	0.7	674	12	AO018046	AO018046 CIT-HSP-2
724	19	0.7	575	12	BF080637	BF080637 231520.MA	c 797	19	0.7	677	12	TA210B090	TA210B090
725	19	0.7	578	12	AO320254	AO320254 RPTC11-10	c 798	19	0.7	681	10	BF446371	BF446371 7P36d01.x
726	19	0.7	579	10	BM263813	BM263813 1g30h04.Y	c 799	19	0.7	686	10	BC910647	BC910647 602808693
727	19	0.7	581	9	BE237348	BE237348 146510.MA	c 800	19	0.7	687	12	A2205048	A2205048 SP_0099.A
728	19	0.7	581	12	BH173547	BH173547 gm_1SB001	c 801	19	0.7	687	12	AG068534	AG068534 Pan trogl
729	19	0.7	581	12	AO426168	AO426168 CITBI-El-	c 802	19	0.7	688	12	A2427766	A2427766 1M0129B01
730	19	0.7	585	9	A1524093	A1524093 CIT-HSP-2	c 803	19	0.7	689	10	BE897430	BE897430 601439341
731	19	0.7	587	12	AO020453	AO020453 CIT-HSP-2	c 804	19	0.7	692	12	AC095520	AC095520 Pan trogl
732	19	0.7	588	9	AL042206	AL042206 DKE2p434G	c 805	19	0.7	693	9	BB522878	BB522878 BB522878
733	19	0.7	588	9	AV720317	AV720317 AV720317	c 806	19	0.7	694	9	AV323038	AV323038 AV323038
734	19	0.7	589	9	AA455657	AA455657 aa22e04.S	c 807	19	0.7	697	9	A1043118	A1043118 uc68h04.x
735	19	0.7	591	10	WB9121	WB9121 zh69b04.sl	c 808	19	0.7	697	12	AG116993	AG116993 Pan trogl
736	19	0.7	591	10	BE285878	BE285878 601108369	c 809	19	0.7	698	9	AM557864	AM557864 L0287F03-
737	19	0.7	591	10	BE377777	BE377777 601229930	c 810	19	0.7	699	12	A2580790	A2580790 1M0369K03
738	19	0.7	592	9	AM556827	AM556827 L0273H07-	c 811	19	0.7	700	12	AO478332	AO478332 RPTC1-11-2
739	19	0.7	593	10	AM141056	AM141056 EST291079	c 812	19	0.7	701	10	BI600186	BI600186 H149E02.E
740	19	0.7	593	10	BE889191	BE889191 601513458	c 813	19	0.7	710	10	BI158942	BI158942 602970482
741	19	0.7	594	9	AV617696	AV617696 AV617696	c 814	19	0.7	711	10	BG400345	BG400345 602464557
742	19	0.7	595	9	AF150340	AF150340 AF150340	c 815	19	0.7	711	10	BI734558	BI734558 603356385
743	19	0.7	600	10	BG806511	BG806511 2031-89.M	c 816	19	0.7	712	10	BJ064032	BJ064032 BJ064032
744	19	0.7	600	10	BG806599	BG806599 2033-03.M	c 817	19	0.7	714	12	AG104913	AG104913 Pan trogl
745	19	0.7	600	10	BI988380	BI988380 3224-56.M	c 818	19	0.7	718	10	BI222242	BI222242 602939719
746	19	0.7	601	9	A1376649	A1376649 te63b06.x	c 819	19	0.7	723	10	A2200561	A2200561 SP_1040.B
747	19	0.7	601	9	AL603336	AL603336 DKE2p686A	c 820	19	0.7	727	10	BE374789	BE374789 601226605

821	19	0.7	728	10	BE369136	BE369136 601222926	894	19	0.7	1179	10	BM479116	BM479116
822	19	0.7	731	9	BS574195	BS574195	895	19	0.7	1188	10	BG166903	BG166903
823	19	0.7	731	12	AG109464	AG109464 Pan trogl	896	19	0.7	1515	11	AK018294	AK018294
824	19	0.7	735	12	BH295208	BH295208 CM230-104	897	19	0.7	1843	10	BF128533	BF128533
825	19	0.7	739	9	BS328667	BS328667	898	19	0.7	2048	11	BC017051	BC017051
826	19	0.7	741	10	BE276357	BE276357 601144005	899	19	0.6	118	10	BE939421	BE939421
827	19	0.7	741	12	BH566840	BH566840 BOGMM47TR	900	18	0.6	132	12	AZ638864	AZ638864
828	19	0.7	745	12	AZ324806	AZ324806 IM0046120	901	18	0.6	123	10	BF090480	BF090480
829	19	0.7	748	10	BG832225	BG832225 602765605	902	18	0.6	141	10	BG007755	BG007755
830	19	0.7	750	12	AZ731909	AZ731909 RPT-24-1	903	18	0.6	141	12	BH010314	BH010314
831	19	0.7	753	9	AU143397	AU143397	904	18	0.6	135	9	AA371471	AA371471
832	19	0.7	754	10	BI970606	BI970606 GM830011A	905	18	0.6	161	9	AM561898	AM561898
833	19	0.7	762	10	BG419708	BG419708 602451939	906	18	0.6	162	10	BF587323	BF587323
834	19	0.7	768	10	BG536481	BG536481 602564582	907	18	0.6	164	9	BA408845	BA408845
835	19	0.7	768	10	BG177764	BG177764 602698053	908	18	0.6	166	9	AI910059	AI910059
836	19	0.7	773	10	BI259848	BI259848 602971507	909	18	0.6	167	12	AZ287256	AZ287256
837	19	0.7	775	10	BI489574	BI489574 603032008	910	18	0.6	169	10	BI022405	BI022405
838	19	0.7	776	10	BI968622	BI968622 GM830005B	911	18	0.6	174	10	BE768656	BE768656
839	19	0.7	787	10	BE887288	BE887288 601510111	912	18	0.6	175	10	BI708541	BI708541
840	19	0.7	791	12	AQ741109	AQ741109 HS-2272-A	913	18	0.6	175	10	BI709210	BI709210
841	19	0.7	792	9	AU130601	AU130601	914	18	0.6	176	10	BI840229	BI840229
842	19	0.7	800	10	BE794809	BE794809 601589269	915	18	0.6	178	9	AA284530	AA284530
843	19	0.7	807	10	BI227029	BI227029 602952184	916	18	0.6	181	9	AA213773	AA213773
844	19	0.7	808	10	BG777192	BG777192 602664414	917	18	0.6	181	12	AZ563534	AZ563534
845	19	0.7	817	10	BF791220	BF791220 602250735	918	18	0.6	195	10	BG218342	BG218342
846	19	0.7	818	9	AL048542	AL048542 DKFZP5860	919	18	0.6	195	12	AZ803588	AZ803588
847	19	0.7	818	10	BI522164	BI522164 603081575	920	18	0.6	197	9	BH256625	BH256625
848	19	0.7	830	10	BG432658	BG432658 602496329	921	18	0.6	200	9	AA333524	AA333524
849	19	0.7	838	10	BI408930	BI408930 602966412	922	18	0.6	202	10	BG202154	BG202154
850	19	0.7	838	10	AK007385	AK007385 Mus muscu	923	18	0.6	203	10	BF411599	BF411599
851	19	0.7	840	9	AU132370	AU132370	924	18	0.6	204	9	AA616935	AA616935
852	19	0.7	845	9	AU130169	AU130169	925	18	0.6	206	10	BG212632	BG212632
853	19	0.7	849	10	BI838890	BI838890 603087272	926	18	0.6	207	10	BF736582	BF736582
854	19	0.7	850	12	AO900025	AO900025 HS-3162-A	927	18	0.6	215	9	AM026744	AM026744
855	19	0.7	853	10	BG924061	BG924061 602822278	928	18	0.6	217	9	BB038977	BB038977
856	19	0.7	859	10	BF673336	BF673336 602136194	929	18	0.6	219	9	AI915039	AI915039
857	19	0.7	862	10	BG683381	BG683381 602651482	930	18	0.6	220	10	BE303653	BE303653
858	19	0.7	867	10	BE253154	BE253154 601116687	931	18	0.6	223	10	BF744093	BF744093
859	19	0.7	868	10	BF684872	BF684872 602142589	932	18	0.6	224	12	AZ222734	AZ222734
860	19	0.7	870	12	CNS03TYO	AL260453 Tetradon	933	18	0.6	227	9	AI962077	AI962077
861	19	0.7	872	9	AL539493	AL539493	934	18	0.6	228	10	BG006987	BG006987
862	19	0.7	872	10	BG771220	BG771220 602721568	935	18	0.6	228	12	AQ342298	AQ342298
863	19	0.7	898	10	BF969103	BF969103 602270057	936	18	0.6	231	12	AQ026938	AQ026938
864	19	0.7	899	10	BI456360	BI456360 603172870	937	18	0.6	233	10	T30346	T30346
865	19	0.7	903	10	BG117671	BG117671 602349630	938	18	0.6	234	9	BB554958	BB554958
866	19	0.7	905	10	BI729498	BI729498 603349142	939	18	0.6	235	10	D20836	D20836
867	19	0.7	922	9	AU132930	AU132930	940	18	0.6	241	9	AV281527	AV281527
868	19	0.7	931	12	CNS06DDT	ALJ39592 T3 end of	941	18	0.6	241	9	BB084006	BB084006
869	19	0.7	932	10	BG259881	BG259881 602371838	942	18	0.6	244	10	T03318	T03318
870	19	0.7	934	10	BF684705	BF684705 602141081	943	18	0.6	244	12	AZ644493	AZ644493
871	19	0.7	945	10	BE973483	BE973483 601652319	944	18	0.6	246	9	AM590258	AM590258
872	19	0.7	947	9	AU067250	AU067250	945	18	0.6	247	9	AA844604	AA844604
873	19	0.7	949	10	BF668815	BF668815 602140781	946	18	0.6	248	9	AI637899	AI637899
874	19	0.7	952	9	BB618414	BB618414	947	18	0.6	252	10	F34684	F34684
875	19	0.7	959	9	AU167354	AU167354	948	18	0.6	253	9	AA076896	AA076896
876	19	0.7	960	10	BM475825	BM475825	949	18	0.6	255	10	BM042601	BM042601
877	19	0.7	972	10	BF976622	BF976622 602244114	950	18	0.6	257	9	AV061040	AV061040
878	19	0.7	975	10	BG115887	BG115887 602317303	951	18	0.6	258	10	BE672134	BE672134
879	19	0.7	979	10	BG576433	BG576433 602595675	952	18	0.6	259	9	AI825519	AI825519
880	19	0.7	980	10	BG117509	BG117509 602347637	953	18	0.6	263	9	AI019402	AI019402
881	19	0.7	991	10	BG242226	BG242226 602334590	954	18	0.6	263	9	AI623173	AI623173
882	19	0.7	991	10	BG683980	BG683980 601436727	955	18	0.6	264	10	BG218979	BG218979
883	19	0.7	992	10	BE893396	BE893396 602436727	956	18	0.6	265	9	AA078649	AA078649
884	19	0.7	996	12	CNS012RR	AL101937 Drosophila	957	18	0.6	265	9	BB004173	BB004173
885	19	0.7	997	10	BG758928	BG758928 602713040	958	18	0.6	265	9	BB382159	BB382159
886	19	0.7	1012	12	BI050312	BI050312	959	18	0.6	269	9	AM196911	AM196911
887	19	0.7	1014	10	BI050312	BI050312	960	18	0.6	269	10	T277722	T277722
888	19	0.7	1020	10	BI759069	BI759069 603042864	961	18	0.6	271	9	BB441707	BB441707
889	19	0.7	1023	10	BG342230	BG342230 602374307	962	18	0.6	272	9	AI956496	AI956496
890	19	0.7	1072	10	BF339506	BF339506 602039163	963	18	0.6	273	9	BB473487	BB473487
891	19	0.7	1084	10	BE562821	BE562821 601336529	964	18	0.6	274	10	BG187038	BG187038
892	19	0.7	1095	10	BE791043	BE791043 601586504	965	18	0.6	274	10	BF998461	BF998461
893	19	0.7	1101	12	CNS05JIT	ALJ340238 Tetradon	966	18	0.6	277	9	BB089729	BB089729

967	18	0.6	278	9	AM304018	xv15h11.x
c 968	18	0.6	278	9	BB564825	BB564825
969	18	0.6	278	10	R10355	R10355
c 970	18	0.6	279	9	AV260992	AV260992
971	18	0.6	279	10	BG211877	BG211877
c 972	18	0.6	280	10	BG193003	BG193003
973	18	0.6	281	9	AV144913	AV144913
c 974	18	0.6	281	9	BB415476	BB415476
c 975	18	0.6	283	9	AI800894	AI800894
c 976	18	0.6	284	9	AA843970	AA843970
c 977	18	0.6	284	9	AA312592	AA312592
c 978	18	0.6	284	10	F03657	F03657
c 979	18	0.6	284	12	A293429	A293429
c 980	18	0.6	284	12	A2575767	AST-T32B0
981	18	0.6	285	9	BB360814	BB360814
c 982	18	0.6	285	10	BE571374	BE571374
983	18	0.6	286	10	BE768710	BE768710
c 984	18	0.6	288	9	AM003468	AM003468
c 985	18	0.6	288	10	BE502516	BE502516
986	18	0.6	289	9	AI652715	h219f09.x
c 987	18	0.6	289	9	AV122743	AV122743
988	18	0.6	290	10	BG203005	BG203005
c 989	18	0.6	291	9	AA334780	AA334780
c 990	18	0.6	292	9	AM669561	AM669561
c 991	18	0.6	293	10	BG208440	BG208440
c 992	18	0.6	296	10	BF931762	BF931762
c 993	18	0.6	298	9	AI475015	AI475015
c 994	18	0.6	298	10	N75923	N75923
995	18	0.6	299	10	Z21091	Z21091
c 996	18	0.6	300	9	AI244521	AI244521
c 997	18	0.6	301	9	BB391078	BB391078
998	18	0.6	302	9	AA833475	uc91b11.r
c 999	18	0.6	302	9	BB116648	BB116648
1000	18	0.6	302	9	BB519406	BB519406

ALIGNMENTS

RESULT 1	BM450709	1082 bp	mRNA	linear	EST 05-FEB-2002
LOCUS	BM450709				
DEFINITION	AGENCOURT_6394734 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494415				
ACCESSION	BM450709				
VERSION	BM450709.1	GI:18499749			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
TITLE	1 (bases 1 to 1082)				
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cs9aps-r@mail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.llnl.gov				
	Plate: L1M12119 row: k column: 24				
	High quality sequence stop: 585.				
FEATURES	Location/Qualifiers				
SOURCE	1..1082				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5494415"				
	/clone_lib="NIH_MGC_67"				
	/tissue_type="retinoblastoma"				

/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 298 a 261 c 292 g 226 t 5 others

ORIGIN

Query Match 25.0%; Score 721; DB 10; Length 1082;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 721; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	849	atcaaatatgaagatttgccagccttatacaaggaagcaaacctttagctccag	908
Db	2	ATCAAAATATGAAGAGCTTTGGCCAGCCTTATCAAGAGATCAAACTCTTAGCTCCAG	61
Qy	909	aagacacaaactgggagacacttatactagctagtgaggagacagcttggcagt	968
Db	62	AAGACACAAACTGGGAGACACTTATGATGTACATGAGTGGGAGACAGCTTGGCAGT	121
Qy	969	atgtcaatcctatacaaaacccaaagacacactctgggggaaagccttatgtgtcagg	1028
Db	122	ATGTCAATCTCTCATCAAAAACCAAGACACACTCTGGGGAAACCTTATGTGTGCAAG	181
Qy	1029	gaatgtggcggcagcctttagcgttggaagtcacacacatcacatcagaagacacatca	1088
Db	182	GAAATGTGGCGCAGGCTTTTACGTGGAAGTCAAACTGATACATCAGACAGACACATCA	241
Qy	1089	ggggagaaaccttatgtgtgcaagattgtgagcagagccttacttggaaagtcgaacctc	1148
Db	242	GGGAGAAACCTTATGTGTGCAAGATGTGTGACGAGCGCTTATGGAAGTCGAACCTC	301
Qy	1149	tttaacatcagcggacacacactcagagcctttagtgttgaaagaaatgtggcag	1208
Db	302	TTTACACATCAGCGGACACACTCAGGCTCAAGCTTATGTGTGCAAGATGTGGCAG	361
Qy	1209	agcctttagccttaagtcacacacttataccacacagagcgccacacttggagagaacct	1268
Db	362	AGCTTATAGCTCATGATCAAACTCAATTAACCAACAGAGGCGCCACACTGGGAGAACT	421
Qy	1269	tatgttttcaggaagatgtggcgtgtgcttgcagacatcaccctgttcagacagaag	1328
Db	422	TATGTTTTGGAGGGAATGTGGGCTTCCCGCATTCACCTGTGTCAGACACAAG	481
Qy	1329	agacacacatcagagagaagccttaccatttcagggagtgtagcagaagccttagcag	1388
Db	482	AGACACATTCAGAGAGAGAGAGCTTACATTTTCAGAGGAGTGTGACCAAGCTTTAGCCAG	541
Qy	1389	aagtcacacatcagacacacttaagagacacacacagagagaagccttatgtatgcaca	1448
Db	542	AAAGTCACACCTCATGACAGACACTTAAGAGACACACAGAGAGAAACCTTATATGACACA	601
Qy	1449	gaatgtggcgttcacttaagctgtgaaatcaaacctcaaaaacaccacagagacacactca	1508
Db	602	GAAATGTGGCGCTTATGTGTGCTGGAATCAAACTCAAAACACACACAGGACACACTCA	661
Qy	1509	gggggttaaaccttatgtctgccttgagtgagtgagtgagtgagtgagtgagtgagtgagtgag	1568
Db	662	GGGGTTAAACCTTATGTGTGCTGCTGGAATCAAACTCAAAACACACACAGGACACACTCA	721
Qy	1569	a 1569	
Db	722	A 722	

RESULT 2	BM470740	1035 bp	mRNA	linear	EST 05-FEB-2002
LOCUS	BM470740				
DEFINITION	AGENCOURT_6475358 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5579278				
ACCESSION	BM470740				
VERSION	BM470740.1	GI:18519782			

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1035)
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgabgs-remail.nih.gov Tissue Procurement: ATCC
COMMENT	CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LHAM12335 row: k column: 23 High quality sequence stop: 653.
FEATURES	Location/Qualifiers 1..1035
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5579278" /clone_id="NIH_MGC_67" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: PCMV-SPOrt6; Site:1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT	259 a 249 c 268 g 255 t 4 others
ORIGIN	
Query Match	24.4%; Score 677; DB 10; Length 1035;
Best Local Similarity	100.0%; Prid. No. 0;
Matches 677; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy 1655	cacgaccagaggaaacactcaggggagaagccattgatgtcgatgtgcgaag 1714
Dd 25	CACGACCAGAGAACACTCAGGGGAGAAGCCATTGTATGTCTGAAGTGACGAG 84
Oy 1715	cttaaigaatagtccacctcatlccaacagagacacatlcaagggaaaagcc 1774
Dd 85	CTTAATGATGATGTCACACCCCTCATTTCAACACAGAGACACTTCAGGGAAAGCCTT 144
Oy 1775	tatgtcagggagtgtgacagaagtttcgcagagaagccctaactcgtlttagcacagaag 1834
Dd 145	TATGTGACAGGGAGTGTGCGAGAAAGTTTCGCGAAGCCTTAACCTGTTAGCGACAAGAG 204
Oy 1835	ggcacactcagtgcttgtgtgtgcaaggagatgtggcaagcccttttggcctaagttaac 1894
Dd 205	GGCACACTCAGGTGCTTTGTGTGACAGGAGTGTGGCAAGGCTTTTGGCTAAGTTAAC 264
Oy 1895	tctcataaacacacagagacacagcgaggggagagccctcgtgtgcaggagatgttg 1954
Dd 265	TCTCATTTAAACACAGAGACAGCACAGCGAGGGGGAAGCCTCATGTGTGACGAGGATGTGG 324
Oy 1955	gcaaagccttagccgacagtcacacctaatatagacacagagagacacatlcaagagaga 2014
Dd 325	GCAAAGCCTTAGCCCGGACGTGACACCTCATTTAGACACAGAGACACATTCAGAGAGAGA 384
Oy 2015	gccctatatattgcagaagaagtgtgacggggccttagtcggaaagtcacaacctataagaca 2074
Dd 385	GCCCTATATTTTTGCCAAAAGTGTGAGCGGGGCTTTAGTCCGAGGTCCAACCTTATAGACA 444
Oy 2075	tcaagagacacactcagagatagaacttatgtfataagggaatgtgtfaagcccttag 2134
Dd 445	TCAAGAGCACACTCAGAGATAGAAACTTTATGTGTATAGGGAATGTGATACAGCCCTTAG 504
Oy 2135	ccaggagtcatacttcatacagacacacagagacacacacagatgcgtgtgtcttttccagcc 2194

[illegible]

OY	1798	ggttcgcgcagagaagccttaacccgttctttaggcacaaagagggcacaaccacagtgctcttgcgt	1857
Db	156	ggtttccgcgcagagaaccttaacctgttttaggcacaaagagggcacaactcagtgctcttgcgt	215
OY	1858	gcagaggagctgtgycgcagagccttctgtctgaagttaactctcataaacaaccagagagcac	1917
Db	216	gcagaggaactgtgtggcccaaaagcctttttgtgcctaaattgaactcctcaattgaacccagagagcac	275
OY	1918	acgcagggggggaagcctatctgtctgcagggagctgtgggcgaagccttaagccgcagctaac	1977
Db	276	acgcaggggggggaagcctctctgtgtcctcagaggaactgtgggcgaagccttttaacccgcagctaac	335
OY	1978	aacctatctagaacaacagagagggacacatctcagagagaagacctatatacttcagagaagtgtc	2037
Db	336	acctctattagacacacacagagagcacacttcagagagagagccttaattatttcgcagaaaagtgcg	395
OY	2038	gacggcgctcttaagctctggagaagctcaacctataagacatacagagagacacactcaagataga	2097
Db	396	gacggcgctcttaagctctggagaagctcaacctattatagacatcaccagagacacactcaagataga	455
OY	2098	aacctatctgtatagaggagaatgtgtgtacagcctttagccagagagatcatactcatcaagac	2157
Db	456	aacctattatgtatagaggagaatgtgtgtacagcctttagccagagagatcatactcatcaagac	515
OY	2158	accagaagagacaacaacagctgtcgtgtgcttttccagaccatctgtatagataccaaagtggaga	2217
Db	516	accagaagagacaacaacagctgtcgtgtgcttttccagaccatctgtatagataccaaagtggaga	575
OY	2218	catctctgtctgtatctatcatcatcgaagactgtactgtgttaaagactgtcatctccatccacctg	2277
Db	576	catctctgtctgtatctatcatcatcgaagactgtactgtgttaaagactgtcatctccatccacctg	635
OY	2278	aagagaatctgtgctcaattcttcagagagccctgcc	2313
Db	636	aagagaatctgtgctcaattcttcagagagccctgcc	671

```

RESULT      4
LOCUS       AM245709
DEFINITION  2833003.Sprline NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2833003 5',
ACCESSION   AM245709
VERSION     AM245709.1
KEYWORDS    GI:5588702
SOURCE      EST.
            human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 590)
            NIH-MGC http://mgc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            other ESTs: 2833003, 3prtime
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: DCTD/DRP cDNA Library Preparation: Ling
            Hong/HuIn Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (ILNLC) DNA sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/ILNLC at:
            www.bio.lnlnl.gov/bdnp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center::
            http://www.genome.washington.edu
            Plate: LICM10 row: O column: 4
            High quality sequence stop: 582.
            Location/Qualifiers
                1..590
                /organism="Homo sapiens"

```

Query Match	Best Local Similarity	21.0%: Score 581; DB %; Length 590;
Matches 581;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY 544	ggcgaggaatccctccacctcgtggaaacacatcccggaagatcagaacaacagcgg	603
Db 10	GGCGGGAATCTCTCCACCTGGGAAAACCTATCCGAGGATCGAAACACGCGKG	69
QY 604	atccattcgtcttagtgcgaagcagaatgatcgaagaggagaaacatccagatcc	663
Db 70	ATCCATCTCGCTTGTGTGCGAAGAGGAAATGATTCAGAGGGAGAAACATCCAGACTCC	129
QY 664	tgtttggagagtaagcaaaaatggcacttcaaaagccattccagccactgaagaac	723
Db 130	TGTTTGGAGAGTAAAGCAAAAATGGCACTTCAAAAGGCACTTCCAGCCCACTCGAAGAAC	189
QY 724	aacagccagcagcagcagcaaggaagacaacacagatggtatagataggtccagccctgaac	783
Db 190	AACAGCCAGCAGCTCCAGGAGAACACACAGTGTGTGATATAGGTCCAGCCTGTAAC	249
QY 784	ggaggccagatcctagaggaacaacagacaagatgatcagtggttagaagttcagatttg	843
Db 250	GGAGGGCAGATCTAGAGGAAACAGCAAAAGTATTCAGTGTGTAAGAACTCAGGATTTG	309
QY 844	gagaaatcaaatatgaagaagtttggcgcaagccttatcaaggaagtcaaacctccatgac	903
Db 310	GAGAAATCAAAATATGAAAGTTTGGGCCAGGCTTTATCAAGAGATCAAACTCTTAGCC	369
QY 904	tcacgaagacaacaactcgtggagacaacctcatgatcagctgagtggtggagacagcttg	963
Db 370	TCCAGAAACACAAACTGGGAGACACTTTCATATACATGATGAGTGGGAGACAGCTTGG	429
QY 964	gcagatgatcagtcctcatcaaaaaaccacaagacaacactctgggggaaagccttatgt	1023
Db 430	GCAGTATCTACGTCTCTCAAAAAACCCAAAGACACACTCTGGGGGAAAGCTTATGTGT	489
QY 1024	gcaggaatctggcggaagcttaccgttggaagtcacaacctgatcacacatcagaagacac	1083
Db 490	GCAGGAGATGTGGCGGAGGCTTTAGCTGGAAGTCAAACTATCATCATCAGAGGACAC	549
QY 1084	actcagggagaaacctatgtgtgcagaagatgtlbgacga	1124
Db 550	ACTCAGGAGAAACCTTATGTGTGCAAGGATTTGGACGA	590

BASE COUNT 192 a 127 c 161 g 110 t

ORIGIN

/db_xref="taxon:3906"
 /clone="IMAGE:2823003"
 /clone_lib="NIH_MGC_7"
 /issue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAC(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

REFERENCE 1 (bases 1 to 1009)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 TITLE NIH-MGC http://mgc.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM12117 row: m column: 14
 High quality sequence stop: 649.

FEATURES
 source
 Location/Qualifiers
 1..1009
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5493685"
 /clone_lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies." 259 g 224 t 2 others

BASE COUNT
 289 a 235 c 259 g 224 t 2 others

Query Match 19.7%; Score 545; DB 10; Length 1009;
 Best Local Similarity 99.8%; Pred. No. 9.4e-276;

Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 206 ggcgttcgtgcatccagagatggtgctgtgacttcaccacagagagatgagttgtt 265
 |||||
 Db 196 ggcgttcgtgcatccagagatggtgctgtgacttcaccacagagagatgagttgtt 255
 OY 266 gggccctgctcagagagaccctgcacagagggtgagtctgagactataaccatctgt 325
 |||||
 Db 256 gggccctgctcagagagaccctgcacagagggtgagtctgagactataaccatctgt 315
 OY 326 ctccactggaattccatctctaaacaaactcattgtctgcgtgagcgaggaggaaac 385
 |||||
 Db 316 ctccactggaattccatctctaaacaaactcattgtctgcgtgagcgaggaggaaac 375
 OY 386 ggcctgagagagagagaaatgtccacttgacacctgtccagaaatcgaaagcagaat 445
 |||||
 Db 376 ggcctgagagagagagaaatgtccacttgacacctgtccagaaatcgaaagcagaat 435
 OY 446 tcaacttagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 505
 |||||
 Db 436 tcaacttagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 495
 OY 506 gtggtctgagatctctctcctcctcctcctcctcctcctcctcctcctcctcctcct 565
 |||||
 Db 496 gtggtctgagatctctctcctcctcctcctcctcctcctcctcctcctcctcctcct 555
 OY 566 gggaaacactatccagaagatccagaacacagacagatccatctgcttagtgagaa 625
 |||||
 Db 556 gggaaacactatccagaagatccagaacacagacagatccatctgcttagtgagaa 615
 OY 626 agcagaatgattcagaag 685
 |||||
 Db 616 accagaaatgattcagaag 675
 OY 686 tggagcttaag 745
 |||||
 Db 676 tggagcttaag 735

OY 746 agacacacagtggtgatataggtccagccctgacagagagagagagagagagagag 801
 |||||
 Db 726 agacacacagtggtgatataggtccagccctgacagagagagagagagagagagag 791

RESULT 6
 A1669533/c 563 bp mRNA linear EST 14-MAY-1999
 LOCUS wb88h09.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2312801.3',
 DEFINITION mRNA sequence.
 ACCESSION A1669533
 VERSION A1669533.1 GI:4834307
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 563)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/bbrp/image/image.html
 Seq primer: -400p from Gibco
 High quality sequence stop: 454.

FEATURES
 source
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="2312801"
 /clone_lib="NCI_CGAP_Pr28"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr28 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
 162 a 125 c 150 g 126 t

Query Match 18.4%; Score 510; DB 9; Length 563;
 Best Local Similarity 99.8%; Pred. No. 2.5e-257;

Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2212 tggagacattctgtgtgatataatgcatgagactgactgtaagactgtatccatc 2271
 |||||
 Db 563 tggagacattctgtgtgatataatgcatgagactgactgtaagactgtatccatc 504
 OY 2272 cacctgaaagagaattgctggtcattttccagagagccctgccccttccactggtgagtg 2331
 |||||
 Db 503 cacctgaaagagaattgctggtcattttccagagagccctgccccttccactggtgagtg 444
 OY 2332 tgggtgtgaaacccggtcaggttaatgagtgtgagagagagagagagagagagagag 2391
 |||||
 Db 443 tgggtgtgaaacccggtcaggttaatgagtgtgagagagagagagagagagagagag 384
 OY 2392 gataaggtgtggttaacctgtgaaccacctaagactgaagagacagtcgcggttaatc 2451

```
|||||
Db 383 GATAGGGGTGCTACCGTGGTGAACCCAAACCTTAAAGCTGAAGACAGTCCGGCTAATC 324
QY 2452 ctaatactgaattgagaacactgctctccatttggtggtgttcctccgattatcccaa 2511
Db 323 CTGATGATGATGATGAGAACTGTCTTCCATTTGGTGTGCTTCCGATGATATCCCAA 264
QY 2512 cccctcaaccattttacgtaacactgctctccattggttttcaactgctgtgcc 2571
Db 263 CCTTACACCTATTTAGTATACCTGCGCTTTCCTAATGTTTATACATGCTGTCGCC 204
QY 2572 accctttgagtggtgctcttgcaacttaacaatcagtcagtgtaattccctattctg 2631
Db 203 ACCTTTGAATGCTGCTGCTTGCATCTACTACAAATCAAGTCAACGTGATTCCTTATTCG 144
QY 2632 agcccaataaagaccagactgactgcaatgagagagaagaatcaaccctgtctgaggt 2691
Db 143 AGCCCATAAAGAACCCAGACTGAGTGAAGAGAAATCACCTGCTGTGGGGGT 84
QY 2692 tggggaaccactccctcactccctcctcactgagagctgtctcttctgcaataaattc 2751
Db 83 TGGGGAACACTCCCTGATCCCTGCTGACTGAGAGCTGTCTTTCCTCAATAAATTC 24
QY 2752 tttcttaaccatacctcaacct 2772
Db 23 TTTTCTAACCATCTCACCCT 3

RESULT 7
AL047439 535 bp mRNA linear EST 29-FEB-2000
LOCUS DKFP586I0320.r1 586 (synonym: hute1) Homo sapiens cDNA clone
DEFINITION DKFP586I0320, mRNA sequence.
ACCESSION AL047439
VERSION AL047439.1 GI:4727354
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Mambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Mambutt, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Mambutt R
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFP586I0320) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..535
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFP586I0320"
/clone_1lb="586 (synonym: hute1)"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT 159 a 128 c 143 g 105 t
ORIGIN
Query Match 18.4%; Score 509; DB 9; Length 535;
Best Local Similarity 100.0%; Pred. No. 8.2e-257;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1071 calcagagacacactcaaggagaaacctatgtgtgcaaggattgtgacagggctt 1130
Db 1 CATCAGAGGACACACATCAGGGGAGAAACCTTATGTGTGCAGAGATTGTGACAGAGCTTT 60
QY 1131 acttgaagtcgaacctcttaccatcagcgagacactcaaggctcaagccttatgtg 1190
Db 61 ACTTGAAGTCGAACCTTTTACATCAGCGGACACACTCAGAGGCTCAAGCCTTATGTG 120
QY 1191 tgaagaagaatgtgggaagcctttagcctgaagtcaacctcatctccacagaaggcg 1250
Db 121 TGCAGGAATGTGGGAGAGACCTTTACCTGAAGTCAAACTCAATTCACCAACGAGAGGGG 180
QY 1251 cacacgggagagaagccttatgtcttcaaggaaatgtggcgctgtcttcgcagcatca 1310
Db 181 CACACTGGGAGAGAGCCTTATGTTTCAGGAGATGTGGCGTGTGCTTCCGACGATTTCA 240
QY 1311 cactgtgtcagacacaagaagacacatcagaagagaagccttatacttgcaaggagtg 1370
Db 241 CACCTGTCTAGACACAAGAGACACATTCAGAGAGAAGCCTTACATTTGACAGGAGTGT 300
QY 1371 gacagaagccttagcagaagtcacacactcatcagaacacttaagagacacagaagag 1430
Db 301 GAGCAAGGCTTTAGCCAGAGATCACACTATCAGACACTTAAGACACACAGAGAGAG 360
QY 1431 aagccttatgtatgcacagaatgtggcgctcactttagctgtgaaatcaaacctcaaa 1490
Db 361 AAGCCTTATGTATGACACAGATGTGGGCGTCATTTAGCTGGAATCAAAACCTCAAAACA 420
QY 1491 caccagaagagacacactcaagggttaaaccttatgtctgctgtgagtgcggaagctt 1550
Db 421 CACCAAGGACACACATCAGAGGGTTAAACCTTATGCTGCTGAGTGGCGGAGTCTTT 480
QY 1551 agcctgaagtcacaaccttaacaacacca 1579
Db 481 AGCCTGAATCAACCTTAAACAAACACCA 509

RESULT 8
AI802142 558 bp mRNA linear EST 16-DEC-1999
LOCUS tx30c05.x1 NCL_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2271080 3',
DEFINITION mRNA sequence.
ACCESSION AI802142
VERSION AI802142.1 GI:5367614
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 558)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA library Preparation: M. Bento Soares, Ph.D.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
www.bio.lnl.gov/bhrp/image/image.html
Insert Length: 771 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 455.
FEATURES
source
1..558
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2271080"
```


QY 2716 ctccactgagagctgtcttcttgcataataaattcttcttaccatccca 2768
|||||
Db 57 CTCACACGAGAGCTGTCTTCTTGTCTCAATAAATCTTTCTACCATCTCA 5

RESULT 10
BE178716 624 bp mRNA linear EST 22-JUN-2000
LOCUS PM4-HT0606-030400-001-cl0 HT0606 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE178716
ACCESSION BE178716.1 GI:8657868
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 624)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FARESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=est&PM4-HT0606-030400-001-cl0&ts=2000-04-03&td=1)
Seq primer: puc 18 forward
High quality sequence start: 9
High quality sequence stop: 603.
Location/Qualifiers
FEATURES
source 1..624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0606"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESNES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 188 a 137 c 177 g 122 t
ORIGIN

Query Match 17.0%; Score 471; DB 9; Length 624;
Best Local Similarity 99.7%; Pred. No. 9.4e-237;
Matches 571; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 734 acagtcacgaagaacacacagtcgtgatagataggtccagccctgaacgagggcaga 733
|||||
Db 9 ACAGTCCACGAGACACACAGTGTGATATAGGCTCCAGCCCTGAAGGAGGGGAGA 68

QY 794 tctcagaagaacagcagaatattcattgcttgaagaagtcaggaatttgagagaatcaa 853
|||||
Db 69 TCTAGAGGAACACGACAAATATTGCTATGCTTTTGAAGTCTCAGGATTTGGAGAAATCAA 128

QY 854 atataagaagtttggccaggcttatcaagagtgcaaacctccttagctccagaagac 913

Db 129 ATATGAGAGCTTGGCGCAGGCTTTATCAAGAGTCAAACTCTTAGCTCCAGGAAGAC 188
|||||
QY 914 acaaacctggggagacacacttaactatgtacactgtgtgggagacagctttggcagatgtc 973
|||||
Db 189 ACAAACTGGGGAGACACCTTACATGATCACTGAGTGGGAGACAGCTTGGCAGTATGTC 248

QY 974 agtctcatcaaaaaaaccaagagacacactctgggggaaacccattatgtgtcagggaatg 1033
|||||
Db 249 AGTCTTATATAAAACCAAGACACACTCTGGGGGAAACCTTATGTGTGCAAGGAGATG 308

QY 1034 tggcgagagcttaccgttgaagtcacacactgtatcacacatcagaagagacacactcaggga 1093
|||||
Db 309 TGGGCGAGGCTTACGCGAAGTCAAACCTGATCACACACAGACAGACACTCAGGGGA 368

QY 1094 gaaacctatgtgtgagaagattgtgagaagagcttacttgaagagcgacactctttac 1153
|||||
Db 369 GAAACCTTATGTGTGCAAGGATTTGTGACGAGGCTTTACTTGAAGTCGAACCTCTTAC 428

QY 1154 acatcagcgagacacactcagagctcactatgtgtcagaagatgtggcagagctt 1213
|||||
Db 429 ACATCAGCGGACACACTCAGGCTCAAGCCTTATGTGTGCAAGGAATGTGGCAGACTT 488

QY 1214 tagcctgaagtcacacacttaaccacagagggcgacacttggggagagacttatgt 1273
|||||
Db 489 TAGCTGTAAGTCAAACTTACCTTACCACAGAGGGCGCACACTGGGGAAGCCTTATGT 548

QY 1274 ttgcaggaagtgtggcggtgtgcttgcagca 1306
|||||
Db 549 TTGCAGGGAATGTGGCGGTGCTTGCAGCA 581

RESULT 11
BF111754/c 512 bp mRNA linear EST 20-OCT-2000
LOCUS BF111754
DEFINITION IMAGE:3523123 3', mRNA sequence.
ACCESSION BF111754
VERSION BF111754.1 GI:10941444
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 447.
Location/Qualifiers
FEATURES
source 1..512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3523123"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pTZ19D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtraacted hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NBHP pool 1: 309384-310919, 323208-325895 Soares NBHP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 -

ACCESSION AM401692
VERSION AM401692.1 GI:6920378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNT at:
www.bio.lnl.gov/dbrip/image/image.html
Seq primer: M13 Forward
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3053549"
/clone_lib="NIH-MGC-36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LT1)"
/note="Vector: pT7m3-Pac; Site.1: NotI; Site.2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonafide, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 138 a 113 c 136 g 98 t 1 others
ORIGIN

Query Match 16.2%; Score 449; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.6e-225;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 ggcagatgtcagtcctccatcaaaacccaagagacacactctgggggaagccttaagt 1022
|||||
DB 8 GGCAGRTATGTGACGTCCATCAAAAACCCCAAGACACACTCTGGGGAAAGCCTTATGTG 67
QY 1023 tgcagggaatgtgagcagccttaagctgaagtcacaactgtatcacatcagaagaca 1082
|||||
DB 68 TGCAGGGAATGTGGGAGGCTTACGTGAGATCAAACTGATACACATCAGAGAGACA 127
QY 1083 cactcagggggaagaacctctgtgtcagaagattgtgagaagccttaacttgaagtcg 1142
|||||
DB 128 CACTCAGGGGAGAAACCTTATGTGTGCAAGATTGTGAGAGAGCTTACTTGAAGTCG 187
QY 1143 aacctcttacatcagcagacacactcagggcctcaagccttatgtgtcagaagaaagt 1202
|||||
DB 188 AACCTCTTTACATCAACGCGACACTCAGGGCTCAAGCCTTATGTGTCAAGGAATGT 247
QY 1203 gggcagagccttaagcctgaagtcacaactcatatccaccagaaggcgacacttgggag 1262
|||||
DB 248 GGGCAGAGCCTTATGACCTGAAGTCAAACTCATTAACCCACAGAGGGGCGACACTGGGAG 307
QY 1263 aagccttaagtcttcagggaagtgtggcgtgtgcttcgcagacatcaacctgtgtcaga 1322
|||||
DB 308 AACCCCTTATGTGTGAGGAGATGTGGCGTGTTCGCCAGCATTCACACCTGTGTCAGA 367
QY 1323 caaagaagacacatcagagagaagccttaacttgcagaggaagtgtgagaagccttt 1382
|||||
DB 368 CACAAAGGACACATTCAGAGAGAGAGCCTTACATTTTGCAGGAGGTGTGAGCAAGGCTTT 427

QY 1383 agccagaagtcacacctcatcagacactt 1411
|||||
DB 428 AGCCAGAGATGCACACCTCATCAGACACTT 456

RESULT 14
LOCUS AV720986 486 bp mRNA linear EST 16-OCT-2000
DEFINITION AV720986 HTB Homo sapiens cDNA clone HTB1C05 5', mRNA sequence.
ACCESSION AV720986
VERSION AV720986.1 GI:10818138
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu
S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
Chen,J., Chen,Z. and Han,Z.
Homo sapiens CDNA HTB clones
Unpublished (2000)
JOURNAL Contact: Zeguang Han
COMMENT Chinese National Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="HTB1C05"
/clone_lib="HTB"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI"

BASE COUNT 104 a 134 c 110 g 137 t 1 others
ORIGIN

Query Match 15.1%; Score 419; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 2.4e-209;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2269 atccacctgaagagaatgtctgtcattctcaaggagcccttcctcactgtgga 2328
|||||
DB 1 ATCCACTGAAGAGAAATTCCTGCTCATTTTACGAGAGCCCTCCTCTCATCTGTGGA 60
QY 2329 tggctgggtgtgtgaaacccggtcaggtaatgatatgtgcagagagagtcacaaagtcgccag 2388
|||||
DB 61 TGGTGGGTTGTGGAAACCCGCTCAGTAAATGATGAGCGAGGAGCAATGATGCCAG 120
QY 2389 gtcagatagggtgtgtgacacctggtgaaacccaacttaagctgaagaacagtcggcgttaa 2448
|||||
DB 121 GCAGATRAGGGGTGGTACCTGGTGAACCCAACTTAAACCTGAAGACAGTCCCGCTGAA 180
QY 2449 atccataactgaatgagaacctgtctccatcttgatgtgtgtcttcctccagatgaccc 2508
|||||
DB 181 ATCCTCATACTGAATGTGAGACCTGCTTCCTCATTTTGGTGTGCTTCCTCCGATGTATCC 240
QY 2509 caacctcaactatttaactgatacttacccttccttaattggttttaactgtgtgt 2568
|||||
DB 241 CAACCCCTACCTAATTTTACGTATACCTGCGCTTCTAATGTTGTTTACACTGCTGTCG 300
QY 2569 ccacacttttgagtggtgtccttgatcacttaacaatcagtcgaagtgatctccctatt 2628
|||||
DB 301 CCCACCTTTTGAAGTGTGCTTGTGATACATTACAAATCACTCAAGCTGATATCCCTATT 360

Qy 2629 ctgagccataaagaccacagctcagctgagagagaatcacccctgctg 2687
|||||
Db 361 ctgagccataaagaccacagctcagctgagagagaatcacccctgctg 419

RESULT 15
AM964787/c 613 bp mRNA linear EST 01-JUN-2000
LOCUS AM964787
DEFINITION ES3376980 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
ACCESSION AM964787
VERSION AM964787.1 GI:8154743
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 613)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@ligr.org
Plate: 207
Seq primer: Forward.

FEATURES

Source 1. 613 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGH"
/note="Vector: pBluescriptSKM"

BASE COUNT 171 a 140 c 158 g 144 t
ORIGIN
Query Match 14.9%; Score 414; DB 9; Length 613;
Best Local Similarity 99.8%; Pred. No. 1.1e-206;
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2223 tctgtgtgattatgcatgagctgtactgtgtaagactgttatccatccactgaaga 2282
|||||
Db 544 tctgtgtgattatgcatgagctgtactgtgtaagactgttatccatccactgaaga 485
Qy 2283 gaattgtgtgcatttttaagagacccctcctcctcactgtgtagtggtgtgta 2342
|||||
Db 484 gaattgtgtgcatttttaagagacccctcctcctcactgtgtagtggtgtgta 425
Qy 2343 aacccggtcaggaatgatagatgagcagagagcagatcaatgccagacagataggggtg 2402
|||||
Db 424 AACCCGCGTGAAGTAAGTAGTGCGAGGAGCAGTCAATGCCAGCAGATAGGGGTG 365
Qy 2403 gtaacctgtgtaaacccaacacttaagactgaagacagctcccgactaaatccatctactga 2462
|||||
Db 364 gtaacctgtgtaaacccaacacttaagactgaagacagctcccgactaaatccatctactga 305
Qy 2463 ttgagaacctgtctccactttgtgtgttctcctcagttgatacccaacccttaacct 2522
|||||
Db 304 ttgagaacctgtctccactttgtgtgttctcctcagttgatacccaacccttaacct 245
Qy 2523 tttaagctacactgtcccttccactaattgttttaactgtctgtcccaacctttgagt 2582
|||||
Db 244 tttaagctacactgtcccttccactaattgttttaactgtctgtcccaacctttgagt 185
Qy 2583 ggtgccttgactactacaatcagtcacagtgtaattccctactcttgagccataaa 2642
|||||
Db 184 ggtgccttgactactacaatcagtcacagtgtaattccctactcttgagccataaa 125

Qy 2643 gaccacagactcagctgcagctgagagagaatcacccctgctg 2687
|||||
Db 124 gaccacagactcagctgcagctgagagagaatcacccctgctg 80

RESULT 16
BE391140 697 bp mRNA linear EST 21-JUL-2000
LOCUS BE391140
DEFINITION 601286705F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3613513 5', mRNA sequence.
ACCESSION BE391140
VERSION BE391140.1 GI:9336505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 697)
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC <http://mhc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM274 row: 1 column: 02
High quality sequence stop: 635.

FEATURES

Source 1. 697 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3613513"
/clone_lib="NIH MGC 44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 185 a 171 c 201 g 140 t
ORIGIN
Query Match 14.9%; Score 413; DB 10; Length 697;
Best Local Similarity 99.6%; Pred. No. 3.8e-206;
Matches 513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1280 ggaattgtggcgtgtgttcgcacagatcacacactgtgtcagaacacaaagagacacatc 1339
|||||
Db 1 ggaattgtggcgtgtgttcgcacagatcacacactgtgtcagaacacaaagagacacatc 60
Qy 1340 agagaagaagccttaacttgacagagagtgtagcaagcctttagcagaatcacacct 1399
|||||
Db 61 agagaagaagccttactatttgacagagagtgtagcaagcctttagcagaatcacacct 120
Qy 1400 catcagaaccttaagagacacacagagagaaagcctttagcagaatgtgtggcg 1459
|||||
Db 121 catcagaacctttaaagacacacacagagagaaagcctttagcagaatgtgtggcg 180
Qy 1460 tcaacttaactggaatacaaacctcaaacacacacagagagaaacactaggggttaaac 1519
|||||
Db 181 tcaacttaactggaatacaaacctcaaacacacacagagagaaacactaggggttaaac 240
Qy 1520 ttatgtctgtcgtgagtggtggtggtcttagcctgaagtcacaaacttaacaaacacca 1579

|||||
Db 241 TTTATGCTGCTGAGTGGCGAGTGTCTTACCTGAACTCAACCTTAACAACACCA 300
Qy 1580 gaagttacacacgggggagaagaccatttattatgtaacgagatgtggcgaagctttaccg 1639
Db 301 GAGGTACACACCGGGGAGAGACCATTTGTATGTACGAGATGGGGGAGGCTTTACCCG 360
Qy 1640 gaaatcaacctgtatcacgacacagagacacactcagggaggaagccatttattgtgc 1699
Db 361 GAAATCAACCCCTGAGCAGCAGACAGACACACTCAGGGGAGAGACCATTTGTATGTGC 420
Qy 1700 tggatgtggaagcggagcttataatgataatccaccctcattcacaccagaagacattc 1759
Db 421 TGAAGTGGACGAGGCTTTAATATGATGTCACCCCTATTTCACACAGAGACACATTC 480
Qy 1760 aggggaaagcccttattgtcagggagtgtagca 1794
Db 481 AGGGGAAAGCCCTTTATGTGTGACGAGAGTGTGCA 515

RESULT 17
BG928720 673 bp mRNA linear EST 06-NOV-2001
LOCUS HNC72-1-B2.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
DEFINITION HNC72-1-B2.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
ACCESSION BG928720
VERSION BG928720.1 GI:14323243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 673)
Kumar,S., Connor,J.R., Dadds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathie,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
IDENTIFICATION and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteoarthritic cartilage cDNA libraries
osteoarthritiis Cartilage 9 (7), 641-653 (2001)
21482651
CONTACT: Sanjay Kumar
UN2109
GlaXosmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@sk.com
Seq primer: 17.
Location/Qualifiers

FEATURES
source 1..673
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HNC (Human Normal Cartilage)"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
Directional"

BASE COUNT 181 a 156 c 189 g 145 t 2 others
ORIGIN

Query Match 14.6%; Score 405; DB 10; Length 673;
Best Local Similarity 99.5%; Pred. No. 6.3e-202;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1655 cagcgcacccagagacacactcagggagagaagccatttattgtctgagtgtagcagag 1714
Db 116 CACGCACCCAGAGACACACTCAGGGGAGAAACCATTTGTATGTGTGAGTGGAGAGAG 175
Qy 1715 cttaataatgtagtccaccctcatttcacacacagagacacattcaggggaaagccctt 1774
Db 176 CTTTATATGATAGTCCACCTTCATTTCACACACAGAGACACATTCAGGGGAAAGCCCTT 235

Qy 1775 tatgtcaggaagtgtgacagaagtttcgagcaagcctaactgtttaagcacaaag 1834
Db 236 TATGTACAGGAGATGTGGCAGAAAGTTTCGGCAGAAAGCTTAACCTTTTAGCACAAAG 295
Qy 1835 ggcacacactaagtgacctgtgtgtcaggaagtgtggaagccttltgttaagtaac 1894
Db 296 GGCACACTAGAGTCCCTTTGTGTGTCAGGAGTGTGGCAAGGCTTTGTGCTTAAGTTAAC 355
Qy 1895 tctcataaaccacaaagacacagcagggaggaagccttaattgttgtcaggaagtgtg 1954
Db 356 TCTCATTAACACCCAGAGAGCACACGAGGGGAAAGCTTCATGTGTGACGGAAGTGTGG 415
Qy 1955 gcaagcctttagccggcagtcacacctcattagaacacagagacacattcaggaagaa 2014
Db 416 GCAGGCTTTAGCCGCGCAGTCACACCTCATTAGACACAGAGACACATTCAGAGAGAA 475
Qy 2015 gccctatattgcagaagtgtgacgggctttagtcggaagttccacctatcagaca 2074
Db 476 GCCTTATATTTTGCAGAAAGTGTGACGNGCTTAACTCCGAAGTCCAACTTATACAGACA 535
Qy 2075 tcaagagacacactcagagtagaacttattgtgtatgaaggaatgtgtacagccttag 2134
Db 536 TCAGAGGACACACTCAGATAGAACTATGTATGAGGAATGTGTACAGCCTTTTAG 595
Qy 2135 ccaggaatcactatcatalcagacacacagagacacacagatgtgtgttlttcaagc 2194
Db 596 CCAGAGTCATPACTTCATCAGACACAGAGACACACAGTGTGTGCTTTTACGCG 655
Qy 2195 attgtcagatcccaagt 2212
Db 656 ATTGCTAGATACCAAGT 673

RESULT 18
AM513018 453 bp mRNA linear EST 03-MAR-2000
LOCUS xt76f03.x1 NCI-CGAP-Ut1 Homo sapiens cDNA clone IMAGE:2792381 3,
DEFINITION mRNA sequence.
ACCESSION AM513018
VERSION AM513018.1 GI:7151096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 453)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
<http://image.llnl.gov/image/html/lifresources.shtml>
Image:llnl.gov/image/html/lifresources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 423.
Location/Qualifiers

FEATURES
source 1..453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2792381"
/clone_lib="NCI-CGAP_Ut1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;

BASE COUNT	129 a	99 c	123 g	102 t
ORIGIN	Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.75 kb. Life Technologies catalog # 11538-014"			

Query Match	14.58;	Score 402;	DB 9;	Length 453;
Best Local Similarity	99.88;	Pred. No. 2.2e-200;		
Matches 452;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps	0;

Qy 2315 ttctctactgtgtagtgggttgtagaaaccggtcagtaatgatgtgycaggaagc 237
|||||
Db 453 TTCTCAGCTGGAGGTGGCTTGTGGAACCCGGTACGTAATGATAGTGGCAGGAGGC 394

Qy 2375 agtcaaatgcccaggcagatagggtgggtactgtgtaaacccaacttaagctgaag 2433
|||||
Db 393 AGTCAATGCCAGGAGATAGGGGTGGGTACTGCTGAACCAACCTTAAGCTGAAG 334

OY 2435 acagtcccgctaatacctcatctactgaattgagaacctgtcttcccaattgtytgcttc 249
|||||
|||
Db 333 ACAGTCCCGCTAAATCCTCTACTACTGAATTGAGAACCCTGTCTTCCCATTTGGTGTCCTTT 274

QY 2495 cctcgcattgatcccaacccctcacccatttaacgtataccctgcctctccctaattgctc 255
|||||
Db 273 cctccgcattgatcccaaccccttcacccattttacgtataacctgcctcttccctaattggtt 214

QY 2555 ttacacactgtgtgccaccctttagtgytgcttgacatactacaatcagtcacg 261
|||||
Db 213 TTTCACACTGCTGTGCCACCCTTTTGAAGTGGTGCTTTCACATACCTACAAATCAGTCAACG 154

QY 2615 tgcattccctattctgagccataaagaccagactcagctgcagtgaagagaat 267
|||||
Db 453 TGTATCCCTATTTCGAGCCATAAAGACCCAGACATCAGCTCAGTGGAGAGAAAT 94
|||||

QY 2675 caccctgctgtgaggttgaggaccatccctgcatccccctccactgagagctgtctc 273
|||||
Db 93 CACCTGCTGTGGGGTTGGGACCACTCCCTCATCCCTCCACTGAGAGCTTCT 34

Qy 2735 ttgcctaataaatctcttctacacatcctc 2767
 |||
 Db 33 TTTGCTCAATAAAATCTTTCTACCATCCGC 1

RESULT 19
AT276016/c

LOCUS	A1276016	481 bp	mRNA	linear	EST 20-NOV
DEFINITION	qw08h02.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:1990515				
ACCESSION	mRNA sequence.				
	A1276016				

VERSION	AI276016.1	GI:3898290
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
MOLLEMAN, H. 1979. *Phylogenie der Säugetiere*. Gustav Fischer Verlag, Stuttgart, 394 pp.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index
Unpublished (1997)
Journal Contact: Robert Strunbaum, Ph.D.
Government

Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCG clone distribution information center
Found at: <http://www.ncbi.nlm.nih.gov/Genbank/Genbank.html>

www-bio.11nl.gov/bbrp/image/image.html
Seq primer: -40UP from G1bco
High quality sequence stop: 414.

FEATURES	Location/Qualifiers
source	1. .481

```

/clone="IMAGE:1990515"
/clone_lib="NCI_GAP-Ut3"
/tissue_type="poorly-differentiated endometrial

```

```

/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.45 kb. Life Technologies catalog #
11541-010"

```

BASE COUNT	129 a	101 c	127 g	124 t
ORIGIN				

Query Match	14.48;	Score 399;	DB 9;	Length 481;
Best Local Similarity	99.88;	Pred. No. 8.7e-159;		
Matches 449; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Oy 2306 gccctgcccttcctcaactgtagatgttggttttggaaaccggtcagtaatatgatgtg 2365
 |||||
Db 481 GCCCGCCCTTCCACACTGAGATGGTGGAATAACC GGTCAGGTAA TGATGATG 422

OY 2366 gcagcagcgacgtccaaatgccagcgcatagggtggtaccgtgtgaaccaactta 2425
|||||
Db 421 GCAGAGCGCATCAATGCCAGGCGCAGATAGGGTGGGTACTCTGGTAACCAACCCTTA 362

QY 2426 aagctgaagacagctcccgctaatactcatactgaattgagaacctgtctccatttg 2485
|||||
Db 361 AAGCTGAAGCAGTCCGCGCTAAATCTCATACTGAATTGAGAACCTGCTTCCCAATTG 302

QY 2486 gtgtgcttcctccgattgatcccaacccttcacctatttttcgataaactgccttcc 2545
|||||
Db 301 GTGTGCTTCTCCCATTTGATCCCAACCTTCACCTATTTTACGATAAAGCTGCCCTTCC 242

QY 2546 taattggttttaacactgcctgtgccacctttgaagfgtgccttgcataacttaacaat 2605
|||||
Db 241 TAAATTGGTTTTTACACTGCTGCTGTGCCACACCTTTGATGTGGTGCCTTTGCATACATTACAAAT 182

Qy 2606 cagttcaacgtgtattccctatctctagagccataaagaccagactcagctcagtgag 2665
|||||
Db 181 CAGTCAACGGTATTCCCTATTCTTAGGCCATAAAGACCCAGACACTCAGTCCAGTGG 122

QY 2666 gagagaatcacccctgctgtgaggttgaggaaacacacccctcctcatccctctccactgag 2725
|||||
|||
Rh 121 GAGAGAAATCATCCCGCTGTGGGGGTTGGGGACACATCCCTGGATCCCTCTCCACTGAG 62
|||||

QY 2726 agctgcttccttgcataataaattcctt 2755
|||||
Dh 61 AGCGTCGTCGTCGTCGTCGTCGTCGTCGTC 32

RESULT 20
AA450184

LOCUS	451 bp	mRNA	linear	EST 04-JUN-19
DEFINITION	zxc42e09.r1 Soares total fetus.Nb2HF8_9w Homo sapiens cDNA clone			
IMAGE:789160 5'	similar to TR:G487783	G487783	ZINC FINGER PROTEIN	
7NE133				mRNA sequence

ACCESSION	AA450184	
VERSION	AA450184.1	GI:2163934
KEYWORDS	EST.	
SOURCE	human	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 451)

AUTHORS
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyl,
Watterson, B. and Wilson, D.

TITLE Washu-Merck EST Project 1997

JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estwats@wustl.edu
 This clone is available royalty-free through LNLN; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -28m3 rev2 ET from Amersham
 High quality sequence stop: 428.

FEATURES
 source Location/Qualifiers
 1. 451

/organism="Homo sapiens"
 /db_xref="GDB:598677"
 /db_xref="taxon:9606"
 /clone_1lb="Soares_Total_fetus_Mb2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7/3D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGGCGCTTAATTTTATTTTATTTT 3']
 (Pharmacia), digested with Not I and Eco RI adaptors
 and Eco RI sites of the modified pT7/3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 122 a 97 c 121 g 111 t

Query Match 14.2%; Score 395; DB 9; Length 451;
 Best Local Similarity 100.0%; Pred. No. 1.1e-196;

Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2020 atatttcgaagaagtgctgacgaggtttagtcgaagtcacacccattatcacagacacaga 2079
 |||||||
 Db 4 ATATTTCAGAAAGTGTGGAGCGGGCTTAGTGGAAAGTCACACCTTATCAGACATCAGA 63
 |||||||
 QY 2080 ggaacactcagagatagaacttattgtataggaatgtgtacagcccttagccagg 2139
 |||||||
 Db 64 GGACACACTCAGATGAGAACTTATGTGTATGGGAGATGTGTACAGCCTTTAGCCAGG 123
 |||||||
 QY 2140 agtcatactcatcacagaccagagagacacacagtcgtgtgctttttcaacattgc 2199
 |||||||
 Db 124 AGTCATCTTCATCAGACACACAGAGACACACAGTGTGTGCTTTTTCAGCCATTTGC 183
 |||||||
 QY 2200 tagataccaaagtggagacatctgtgtgtatcatcatgatacgtactgttaagact 2259
 |||||||
 Db 184 TAAATACCAAAAGTGGAGACTTCTGTGTGATTATGACATGACATGTAAGTACT 243
 |||||||
 QY 2260 tttatctcatcacactgaaagaattgtctgtgtcaatttcagagagcccttccct 2319
 |||||||
 Db 244 TGTATCTTCATCACCCTGGAAGGAAATGTGTGCTCATTTTTCAGAGCCCTTGCCCTTCC 303
 |||||||
 QY 2320 caactgtgagttgtgtgtgtgtgtgaaacccggtcagtgatagatgtgtgagagagcagttca 2379
 |||||||
 Db 304 CAGTGTGAGAGTGTGTGTGGAAACCCGGTCAAGTAATCATATGTGGCAGAGAGGACAGTCA 363
 |||||||
 QY 2380 aatgcccagagcagatagaggtgtgtacctgtgtgaa 2414
 |||||||
 Db 364 AATGCCAGGACAGATAGGGGTGTGTGCTGTGTA 398
 |||||||

RESULT 21
 AI081665 435 bp mRNA linear EST 27-AUG-1998
 LOCUS AI081665/c
 DEFINITION ou69e06.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632514 3',
 mRNA sequence.
 ACCESSION AI081665

VERSION AI081665.1 GI:3418457
 EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA library preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 www.bio.lnl.gov/bnfp/image/image.html
 Insert length: 862 Std Error: 0.00
 Seq primer: 40m3 fwd. ET from Amersham
 High quality sequence stop: 410.

FEATURES
 source Location/Qualifiers
 1. 435

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1632514"
 /clone_1lb="NCI_CGAP_Br2"
 /sex="female, pooled"
 /tissue_type="breast"
 /lab_host="DH10B"
 /note="Vector: pT7/3D-Pac (Pharmacia) with a modified
 polylinker. 1st strand cDNA was prepared from pooled bulk
 breast tumor tissue, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and Eco RI
 into the Not I and Eco RI sites of the modified pT7/3
 vector. This library is the normalized version of
 NCI CGAP Br1.1. Library was constructed by Bento Soares
 and M. Fatima Bonaldo."

BASE COUNT 122 a 90 c 122 g 101 t

Query Match 13.8%; Score 383; DB 9; Length 435;
 Best Local Similarity 99.8%; Pred. No. 2.3e-190;

Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2339 tggaaacccggtcaggttaagtgtgagagagcagtcacaaatgccagagcagataggg 2398
 |||||||
 Db 435 TGGAAACCCGGTCAAGTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 376
 |||||||
 QY 2399 gtgggtacctgtgaaacccaacttaagctgaagacagtcgccgataatccatc 2458
 |||||||
 Db 375 GTGGGTACTGTGGAAACCAACCTTAAGCTGAAACAGATGCCGCTAAATCCTCATAC 316
 |||||||
 QY 2459 tgaatgaaacactgtctcccaattgtgtgtgtcttcccgatagatcccaacccctta 2518
 |||||||
 Db 315 TGAATTGAAGAACCTGTCTTCCATTGTTGTGTGCTTCTCCATGATGATCCAAACCTTTCA 256
 |||||||
 QY 2519 ccattttacgtatatactcccttcctcctaattgtgttttacaactgtgtgccacattt 2578
 |||||||
 Db 255 CCTATTTCGTRTACTCTGCTTCTTCTTAATTGTTTACACTGTGTGCGCCACTTTT 196
 |||||||
 QY 2579 ggtgtgtccttctgcatctacaaatcagtcacagtgatccctatccctatctgagccat 2638
 |||||||
 Db 195 GAGTGTGCTTGTGATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 136
 |||||||
 QY 2639 aaaaagaccagactcagctgcaggtgagagagaaatcacctcgtgtgtgaggttggagac 2698
 |||||||
 Db 135 AAAAGACCCAGACTCAGTGCAGTGAGAGAGAAATCACCTGTGTGTGGGTTGGGGAC 76
 |||||||

Qy	2699	caccctccgataccctccacacgaagagcgctctcttgctcaataaatactcttcta	2758		
Db	75	CACCTCCCTGCATCCCTCTCCACGACGACTGTTCTTTGCTCATTAATAATTTCTTTCTA	16		
Qy	2759	cccatccctaccct 2772			
Db	15	CCCATCCCTCACCT 2			
RESULT	22				
AI750087/c		449 bp	mRNA	linear	EST 22-JUN-1999
LOCUS					
DEFINITION		at35f07.x1 Barstead colon HPLRB7		Homo sapiens	CDNA clone
ACCESSION		IMAGE:2374021 3'			
VERSION		AI750087			
KEYWORDS		AI750087.1 GI:5128351			
SOURCE		EST.			
ORGANISM		human.			
		Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE		1 (bases 1 to 449)			
AUTHORS		Hallier,D., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,			
		Krizman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Marra,M., Martin			
		,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,			
		White,Y., Wylie,T., Waterston,R. and Wilson,R.			
TITLE		WASHU-NCI human EST project			
JOURNAL		Unpublished (1997)			
COMMENT		Contact: Wilson RK			
		Washington University School of Medicine			
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
		Tel: 314 286 1800			
		Fax: 314 286 1810			
		Email: est@watson.wustl.edu			
		This clone is available royalty-free through LIND; contact the			
		IMAGE Consortium (info@image.llnl.gov) for further information.			
		Seq primer: -40UP from GIDCO.			
FEATURES					
source		location/qualifiers			
		1..449			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:2374021"			
		/clone_lib="Barstead colon HPLRB7"			
		/sex="male"			
		/dev_stage="adult, age 25"			
		/lab_host="DH10B (phage resistant)"			
		/note="Organ: colon; Vector: pT7AD-Pac (Pharmacia) with a			
		modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st			
		strand cDNA was primed with a Not I - oligo(dT) primer [5'			
		TGTACGACATCGTAGTGGAGCGCGCCGCTTTTCTTTTCTTTTCTTTT			
		3']; double-stranded cDNA was ligated to Eco RI adaptors			
		[5' ATTACTAGTAAT 3' and 5' ATTACTAGG 3']; digested			
		with Not I and cloned into the Not I and Eco RI sites of			
		the modified pT73 vector. Library constructed by Bob			
		Barstead."			
BASE COUNT		125 a	99 c	119 g	106 t
ORIGIN					
Query Match		13.8%;	Score 383;	DB 9;	Length 449;
Best Local Similarity		99.8%;	Pred. No.2.4e-190;		
Matches 433;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	2322	ctgttgatggtggtgtgtgtggaaccccggttcaggtatgatatagtcgaggaagcgatcaaa	2381		
Db	449	CTGTGTGATGTTGGCTTTGTGAAACCCGGTTCACGTATGATGATGTCAGAGGACACTCAA	390		
Qy	2382	tgcccaagcagataaggggttggttacctgtgtgaaccaaaccttaagctgaagaacatcc	2441		
Db	389	TGCCACGCGCAGATTAGGGGTGGGTACTGCTGTGAACCCACCTTAAGCTGAAGACAGTCC	330		
Qy	2442	cggctaatacctcatalactgaatgtgaacccgtcttcccatcttggttgcttcttcgga	2501		

Query Match 13.6%; Score 378; DB 10; Length 455;
Best Local Similarity 99.8%; Pred. No. 1e-187;
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2341 gaaccgcgtcgaatgatgtatgacagagcagtcacaaatgccagcagatagggt 2400
|||||
Db 429 GAACCCGGTCAGTATGATGTCAGAGGACAGCAAAATCCAGGAGATAGGGGT 370
|||||
QY 2401 gggctacgtggaacccaaccttaagctgaagacagtcgcggcctaactccatctg 2460
|||||
Db 369 GGGTACTGGTGAACCCAACTTAAAGCTGAAGACAGTCCGGCTTAATCTCATACTG 310
|||||
QY 2461 aattgaagaacctgctcccatctggtgtgtcttcctccgattgataccaacctcacc 2520
|||||
Db 309 AATTGAGAAGCTGTCTTCCCATTTGGTGTCCTTCCCTCCGATTGATGACCAACCTTCAC 250
|||||
QY 2521 tatttaacgtataccgtcccttcctccatctgttcttaacacgtcgtgtgccacctttga 2580
|||||
Db 249 TATTTTACGTATACGTCCCTTTCCTAATTGGTTTACACTGCTGTGCCACCTTTTGA 190
|||||
QY 2581 gtgtgctcttgatacttaacaatcagtcagtgatctccctattcttgagccataa 2640
|||||
Db 189 GTGTGCTCTTGCTGACTTCAAAATCAGTCAAGCTGATTCCTTATTTGAGCCCTATA 130
|||||
QY 2641 aagaccagaactcagctcagtgagagagaatcaacctgtctgtgaggttgaggacca 2700
|||||
Db 129 AAGACCAGACTCAGCTGAGTGAGAGAGAAATCAACCCGTGTGGGGGTGGGGACCA 70
|||||
QY 2701 ctccctgcacccctctccacgtgagagctgtcttcttgccaataaattcttctac 2760
|||||
Db 69 CTCCTGCAATCCCTCTCCACAGAGCTGTCTTTTGTCAATAAATCTTTTCTAC 10
|||||
QY 2761 catccctac 2769
|||||
Db 9 CATCTCTAC 1

RESULT 24
BE746403 728 bp mRNA linear EST 15-SEP-2000
LOCUS 601579525F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3528207 5',
DEFINITION mRNA sequence.
ACCESSION BE746403
VERSION BE746403.1 GI:10160395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 728)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DPF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LMCW759 row: a column: 08
High quality sequence stop: 726.
Location/Qualifiers
1..728
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3528207"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
BcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 217 a 169 c 192 g 150 t
ORIGIN

Query Match 13.6%; Score 378; DB 10; Length 728;
Best Local Similarity 100.0%; Pred. No. 1e-187;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1086 tcagggagaacactatctgtgcaaggtttgtgacagagcttacttggaagtcgaac 1145
|||||
Db 244 TCAGGGAGAGAACCTTATGTGTCAAGAGATGTGACAGAGCTTACTTGGAGTGGAC 303
|||||
QY 1146 ctcttacaatcagcagcacactcaggtctcaagccttatgtgtgcaagaaatgtg 1205
|||||
Db 304 CTCTTACACATCAGCGGACACACTCAGGGCTCAACCTTATGTGTGCAAGAAATGGG 363
|||||
QY 1206 cagaacttaagcctgaagcaacctcattaccacagaaggcgacactgggagaag 1265
|||||
Db 364 CAGAGCTTTAGCCTGAAGTCAAACTCATTTACCCACAGAGGCGCACACTGGGGAAG 423
|||||
QY 1266 ccttatgtttgcaaggaatgtggtgtcttgcgcaacttcaacctgtgtcaaac 1325
|||||
Db 424 CCTTATGTGTGACGAGAAATGTGGGCTGTTCGACACTTCAACCTGTGTCAACAC 483
|||||
QY 1326 aagagacacatcagaagagaagccttaccatttgcaaggaatgtgaaagccttaag 1385
|||||
Db 484 AAGAGACACATTCAGAGAGAAAGCTTACATTTGCAAGGAGGTGACGAAGCTTTTAC 543
|||||
QY 1386 cagaagtcacacctcactcagacacttaagagacacacagaagagacactatgtac 1445
|||||
Db 544 CAGAACTCACACCTCATCTCAGACACTTAAGACACACAGAGAGAAAGCTTATGTATGC 603
|||||
QY 1446 acagaatgtggtgtgtac 1463
|||||
Db 604 ACAGAAATGTGGGCGTAC 621

RESULT 25
A1016684 424 bp mRNA linear EST 27-AUG-1998
LOCUS A1016684/C
DEFINITION ou96d05.x1 NCI-CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1635657 3',
DEFINITION mRNA sequence.
ACCESSION A1016684
VERSION A1016684.1 GI:3231020
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 424)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.blo.llnl.gov/bhrp/image/image.html
Insert Length: 1136 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 392.

Source	Organism	Reference	Authors	Title	Journal	Comment
human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 458)	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	Unpublished (1997)
Contact: Robert Strausberg, Ph.D.	Email: cgaps-remail.nih.gov	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.	CDNA Library Preparation: M. Bento Soares, Ph.D.	DNA Arrayed by: Greg Lennon, Ph.D.	DNA Sequencing by: Washington University Genome Sequencing Center	Clone distribution: NCI-CGAP clone distribution Information can be found through the I.M.A.G.E. Consortium/LNLN at: www.bio.lnl.gov/dbip/image/image.html
Seq primer: -40UP from Gidco	High quality sequence stop: 447.	Location/Qualifiers	1..458	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="IMAGE:2472076"
/clone.lib="NCI CGAP_Kid11"	/lab_host="DH10B"	/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids: 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subclaction by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT	129 a	99 c	125 g	105 t		
ORIGIN						
Query Match	13.4%;	Score 372;	DB 9;	Length 458;		
Best Local Similarity	100.0%;	Pred. No. 1.5e-164;				
Matches 372;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY 2316	tcctcactgctgagatggtgtgtgtgtgtgaaacccgagcagtaatgatagtgacgagagca	2375				
DB 458	tcctcactgctgagatggtgtgtgtgtgtgaaacccgagcagtaatgatagtgacgagagca	399				
QY 2376	gtcdaatgccagcagcagataaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2435				
DB 398	gtcaaatgccagcagcagataaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	339				
QY 2436	cagtcgcgcgcgaatcccatcactgacatgtgagaaaccttcctccatctgtgtgtgtgtgt	2495				
DB 338	cagtcgcgcgcgaatcccatcactgacatgtgagaaaccttcctccatctgtgtgtgtgtgt	279				
QY 2496	ctccgattgatacccaaccttcacatattttagatatactgacatgaccttcctcaattgtgtt	2555				
DB 278	ctccgattgatacccaaccttcacatattttagatatactgacatgaccttcctcaattgtgtt	219				
QY 2556	tttaactgt	2615				
DB 218	tttaactgt	159				
QY 2616	gtattcccatcttgagcccatataaagaccagacatcagctcagctgagtagagagaaatc	2675				
DB 158	gtattcccatcttgagcccatataaagaccagacatcagctcagctgagtagagagaaatc	99				
QY 2676	accctgt					
DB 98	accctgt	87				

FEATURES	source
LOCUS	AM140124/c
DEFINITION	501 bp mRNA linear EST 30-OCT-1999
ACCESSION	U7-H-B11.aef-f-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
VERSION	IMAGE:2719247 3, mRNA sequence.
KEYWORDS	AM140124
SOURCE	AM140124.1 GI:6144842
ORGANISM	EST.
REFERENCE	human.
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1 (bases 1 to 501)
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.
	Email: cgaps-remail.nih.gov
	The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library preparation: M.B. Soares lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrrp/image/image.html The following repetitive elements were found in this cDNA sequence: 58-184, >MER52unknown 348-425, >MER52unknown
	Seq primer: M13 Forward
	POLYA=yes.
	Location/Qualifiers
	1. 501
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:2719247"
	/clone_lib="NCI_CGAP_Sub3"
	/lab_host="DH10B (Life Technologies)"
	/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI-CGAP Sub3 library is a subtracted library derived from the NCI-CGAP Sub1 library, which is a subtracted library derived from Bi. Bi constitutes a mixture of 21 normalized or subtracted NCI-CGAP libraries: NCI_CGAP_CO4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_CO10, NCI_CGAP_CO16, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_CO8, NCI_CGAP_CLL1, NCI_CGAP_Le12, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
	NCI_CGAP Kid3 pool 1 LAM 3334-3337, 3682-3683.
	3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 132912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1509004-1522439); NCI_CGAP_GC4 pool 1 LAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_CO10 pool 1 LAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.
	TAG LIB=NCI_CGAP_Kid5

	123	135	106	136	others
BASE COUNT	123	135	106	136	1
ORIGIN	t	g	c	a	

Query Match	13.2%	Score 366	DB 9	length 501
Best Local Similarity	99.6%	Pred. No.	2.2e-181	
Matches 466	Conservative 0	Mismatches 2	Indels 0	Gaps 0

[illegible]

Oy 2360 atagtcgacgaaggcaatcaaatgcccgacagatgggttggtacctgtgtaaaccca 2419
|||||
Db 428 ATAGTGCACGAGCAGTCAAATGCCCGACAGATGGGTGGTACTGCTGAACCCA 369

Oy	2420	accttaagctgtagacagatcccggtcaatcctcatatactgaattggaaaccgctctcc	2479
Dd	368	acctTAAGCTGAAGACAGTCCCGGCTAAATCCTCACTACTGAATTGAGAACCCTGTCTTC	309

Qy 2480 catltggtgcttcctcctcgattgatacccaacccttcaacctatlttaagtataactgc 2539
|||
Db 308 cATTGGTGTCTTCTCCTCGATTGATCCCAACCCCTTNAACCTATTTAAGTATTAACCTGCC 249

Qy 2540 ctttcctaattggttlttaactgcgtgccacccttttgatgtgccttgcataact 2559
 |||||
Db 248 ctTTCCTAATTGGTTTTACACTGCTGTGCCACCPTTTGAGTGCTTGCCTTCATACAT 189

Qy 260 acaaatcatcaagtgatattccctatttcgagccataaagaccagacttcagctgc 2655
 |||||
 Db 188 ACAATCACTCAACGTGTATTCCCTATTTCGAGCCCATAAAGACCACGACTCAGCTGC 129

```

QY * 2660 agtggagagagaatacaccctgtctgtgaggttgggaaccactccctgcacatccccctcc 2712
      |||||
Db 128 AGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGGACCACTCCCTGCATCCCCCTTCC 69

```

Qy 2720 acgagagctgtctcttctgtcaataaatcttctaccatccgc 2767
 |||||
 Db 68 ACTGAGAGCTGTTCTTTGCTCAATTAATTTCTTTCTACCATTCCTC 21

RESULT 29
AT828662/c
AT828662
426 bp
MBDA
116000
ECM 26-ATC

DEFINITION	rrc6f04.xl NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2217535
ACCESSION	mRNA sequence.
VERSION	A1828662
GI	5440333
FASTA	AT828662.1

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Phylum: Metazoa: Chordata: Craniata: Vertebrata: Eutolostom

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 426)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
	National Cancer Institute Cancer Genome Analysis Project (CCAR)	

JOURNAL COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: crabs-rr@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrived by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bhrr/image/image.html

FEATURES

seq primer: -400P from G1bco

High quality sequence stop: 411.

Location/Qualifiers

source

source

1. 426

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2217535"

/clone_1bp="NCI CGAP_Ov23"

/tissue_type="tumor, 5 pooled (see description)"

/lab_host="DH10B"

/note="Organ: ovary; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, Papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"

ORIGIN

Best Local Similarity 99.8%; Pred. No. 8.3e-180;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 426 GTTCAGTATGATAGTGGCAGAGCGCAATCAATGCCAGGCAGATAGGGGTGGGTACT 367

366 GGCGAACCACCTTAAGCTGAGACAGTCCCGGGCTAAATCTCTACTGATTGGA 307

Db 306 ACCTGCTCTCCCATTTGGTGTGCTTTCCCTCCGATGTATGCCAACCCCTTCACCATTTTAC 247

Db 246 GATACCGCCCTTTCCCTAATGGTTTTTACACTGCTGTGCCACCTTTGAGTGGTCCC 187

Db 186 TTTCGACTACTTACCAATATCAGTCAACGTTATTTCCTTCTAGGCCCATTAAGAAGCCCA 127

Db 126 GACTCAGCTGCAGTGAAGAGAAATCAACCCCTGCTGTGGGGGGTTGGGACCACTCCCTGCG 67

Db 66 ATCCCCCTCCACTGAGAGCTGTTCTTTGGCTCATAATAATCTTTTACCCA 13

RESULT	30
A1650983/c	
LOCUS	
DESCRIPTION	
A1650983	
00654927.v1	
NCI CCND CGE	486 bp
HBB8	mRNA
CCND1	linear
EST	04-MAY-1999
IMAGE	2204013_27

ACCESSION mRNA sequence.
 A1650983
 A1650983.1 GI:4734962
 KEYWORDS
 RSCW

SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo

REFERENCE
AUTHORS
TITLE
Tumor Gene Index
1 (bases 1 to 486)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Journal
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue procurement: Christopher A. Moskaluk, M.D.
Ph.D. Michael

R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/db/seq/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 469.
Location/Qualifiers

FEATURES
source

1. 486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2304013"
/clone_lib="NCI-CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonesDB3,
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 136 a 108 c 133 g 109 t
ORIGIN

Query Match 12.8%; Score 355; DB 9; Length 486;
Best Local Similarity 99.8%; Pred. No. 1.4e-175;
Matches 405; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2282 agaatgctggcatttcagagagcccttcctcactgtagtggtagtgg 2341
DB 486 AGAATTGCGGCATTTTCAGAGAGCCCTGCTTCTCCTGATGCTGGTTG 427
OY 2342 aaaccggtcagtgtaagtgcagagagcagtaaatgcccagagatgg 2401
DB 426 AAACCGGTCAGGTAATGATGAGGAGGAGGCAATGCCAGATAGGGGG 367
OY 2402 gttactgtgtgaaccacactaaagctggaagcagtcctggcctaact 2461
DB 366 GGATCTGTGTAAGCAACCTTAAGCTGAAGACAGTCCGGCTTAATCT 307
OY 2462 attgaacactgcttcctcattgtgtgtcttcctcagatgataccact 2521
DB 306 ATTGAACACTGCTTCCATTTGGTGTGCTTCTCCATTTATCTCAACCT 247
OY 2522 atttaacgtacgtcccttcctcactgattttacactgctgtgccac 2581
DB 246 ATTTTACGATACCTGCTTCTTCTTAATGTTTAACTGCTGGGCCAC 187
OY 2582 tgggtccttgatcactcaaatcagtcgaagtgtatccctattctgag 2641
DB 186 TGGTGCTTTGCACTTCAATCAATCAGTCAACCTGATTTCCCTATTCT 127
OY 2642 agaacagactcagctgcagtgagagaaatcacctgcgtctgg 2687
DB 126 AAGACCACTCAGCTCAGTGAAGAGAGAAATCACCTGCTGTGG 81

RESULT 31
BF476173 473 bp mRNA linear EST 05-DEC-2000
LOCUS BF476173/c
DEFINITION naa29c11.x1 NCI-CGAP_Kid11 Homo sapiens CDNA clone IMAGE:3257852
3', mRNA sequence.

ACCESSION BF476173
VERSION BF476173.1 GI:11546989
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 473)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.lnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 464.
Location/Qualifiers

FEATURES
source

1. 473
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3257852"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
plasmid DNA from the normalized library NCI-CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonesDB3 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo.

BASE COUNT 133 a 104 c 126 g 110 t
ORIGIN

Query Match 12.7%; Score 351; DB 10; Length 473;
Best Local Similarity 99.6%; Pred. No. 1.8e-175;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2310 tgccttcctcactgtagtggtagtggtagtggtagtggtagtggtagtgg 2369
DB 458 TGCCTTCTCCTCACTGATGATGGGTTGTGAACCCGTCAGATAGTAGTGG 399
OY 2370 gaggcagtaaatgcccagcagatagggtgtgtactgtgtgaaaccca 2429
DB 398 GAGGCGATGTAATGCCACAGAGATAGGGGTGATCCTGTGAACCCCA 339
OY 2430 tgaagcagtcctcagtaactcactgtaattgagaacctgtcttcctc 2489
DB 338 TGAAGCACTCCGGCTTAATCTCTACTGATGAGAACCTGCTTCCATTG 279
OY 2490 gcttcctcagattgataccaaaccttcacatttlaaglaatactgccc 2549
DB 278 GCTTCCCTCGAATGATCCCAACCTTCACATTTTACGTACCTGCTTCT 219
OY 2550 tgggtttaaactgctgtgccacactttgagtggtgctttgacactta 2609
DB 218 TGGTTTTTACACTGCTGTGCCACCTTTGAGTGTGGCTTGCATTAATCA 159
OY 2610 caacgtgataccctcattctgaagccataaagaccagactcagctgag 2669
DB 158 CAACGTAATTCCTCATTTCTAGAGCCATAAAGACCCAGACTGACGTG 99
OY 2670 gaataccctgctgtgagagttggagacactccctgacactccctcact 2729
DB 98 GAATACCCCTGCTGTGGGGGTTGGGAGACACTCCCTGATCCCTCTCA 39
OY 2730 gtctcttgctcaataaattctttctaccca 2762
DB 38 GTTCTTGTCTCAATTAATTTCTTTCTACCA 6

RESULT 32
AA218658 385 bp mRNA linear EST 12-MAR-1998
LOCUS z936904.f1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:649878 5' similar to gb:M29580 ZINC FINGER PROTEIN
7 (HUMAN);, mRNA sequence.
ACCESSION AA218658
VERSION AA218658.1 GI:1832766
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 385)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 2813 Std Error: 0.00
Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES
source
1..385
/organism="Homo sapiens"
/db_xref="GDB:5276467"
/db_xref="taxon:9606"
/clone="IMAGE:649878"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: plasmid; Site-1:
EcoRI; Site-2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb.
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGACGACGAG
3' -3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT 89 a 102 c 101 g 93 t
ORIGIN

Query Match 12.6%; Score 350; DB 9; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.9e-173;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 ggcgtcggcattgaaggatgctgtgtctaccacccaagaagatggaggtgttg 266
DB 36 ggcgtcggcattgaaggatgctgtgtctaccacccaagaagatggaggtgttg 95
QY 267 agccctgcacagagaccctgcacagaggagtgatctggaactatatacatctgtc 326
DB 96 agccctgcacagagaccctgcacagaggagtgatctggaactatatacatctgtc 155
QY 327 tcaatgaaatcattctcttaacaaaatcattgtctcagctggagcgaggagcg 386
DB 156 TCACGGGAATTCATCTCTTAACCAAAATCATTTGCTCACCTGAGCGAGCGAGCG 215
QY 387 cccttgaagaagaagaagaatgctcactggaacctctgtcgaagaatcgaagaatc 446
DB 216 CCCTGAGAGAGAGAGAAATGTCACCTGACCTTGTCCAGAAATCGAAGCAAAAT 275
QY 447 caactagtcctcctcctcctgatttctccagtcagcaagctcagccaacatgtg 506

DB 276 CAACCTAGTCCCTCTGCTCTGATTTCTCCAGTCAGCAGACTCAGCCACATGTG 335
QY 507 tggcgtgacatctctcaccagctgttccagttatgaggaagcaacc 556
DB 336 TGGCTGAGTCATCTCTCAGCTGTTTCAAGTTATGAGGAGAAATCC 385

RESULT 33
A1986444/c 500 bp mRNA linear EST 20-OCT-2000
LOCUS w265d09.x1 NCI-CGAP_Mel15 Homo sapiens cDNA clone IMAGE:2562929 3',
DEFINITION mRNA sequence.
ACCESSION A1986444
VERSION A1986444.1 GI:5813721
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Chris Moskalk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Gene distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILM at:
www.bio.lnl.gov/db/ftp/image/image.html
Insert Length: 972 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 403.

FEATURES
source
1..500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2562929"
/clone_lib="NCI CGAP_Mel15"
/tissue_type="malignant melanoma, metastatic to lymph
node"
/lab_host="DH10B"
/note="Organ: skin; Vector: pCMV-SPORT6; Site-1: SalI;
Site-2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."

BASE COUNT 139 a 105 c 132 g 124 t
ORIGIN

Query Match 12.6%; Score 350; DB 9; Length 500;
Best Local Similarity 99.8%; Pred. No. 6.2e-173;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2287 tgcctgctacatttcagaagccctccctcctcactgtgagtgtgaaacc 2346
DB 500 TCCTGCTCATTTTCAAGGAGCCCTGCTCTCCTCACTGTGATGGTGTGGAATCC 441
QY 2347 cggtagagtaatgtagtgagcagagcgatcaaatgcccagcagataggggtgtac 2406
DB 440 CGGTCAAGTAATGATGATGCGACGAGGAGCTGAATATGCCAGGACATAGGGGTG 381
QY 2407 ctggtgaaccacaccltaaagctgaagacagtcctcgcttaaatcctcactgaatga 2466
DB 380 CTGGTGAATCCACACTTAAGCTGAAGACAGTCCCGCTAAATCCTACTGAATGA 321
QY 2467 gaacctgtctcccatctgtgtgtcttccctcagatgtatcccaaccttaaccttt 2526
DB 320 GAACCTGCTTCCCATTTGCTGCTTCCCGATGTGATGCCAACCCCTTCATTTT 261

QY 2527 acgtataccctgccccttcccaattgttttacaactgctgtcccaacctttagtgg 2586
|||||
Db 260 AGCTATACCGCCCTTCCATTTGTTTACAGTCTGTGCCACCTTTAGAGTG 201
QY 2587 ccttgcatcactacaataagtcacaglytaltccctcattcttgagccataaagacc 2646
|||||
Db 200 CCTTTCATCTACTTCAAAATGATCAACAGTGTATTTCCCTATTTTGAGCCATTAAGACC 141
QY 2647 caactcagctgagtgagagagaataacacctgctgtg 2687
|||||
Db 140 CAGACTCAGCTGAGTGAGAGAAATCACCTGCTGTG 100

RESULT 34
LOCUS BE395838 551 bp mRNA linear EST 21-JUL-2000
DEFINITION 601310077F1 NIH_MGC_44 Homo sapiens CDNA clone IMAGE:3631603 5',
mRNA sequence.
ACCESSION BE395838
VERSION BE395838.1 GI:9341203
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC <http://imgc.ncl.nih.gov/>
1 (bases 1 to 551)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM321 row: j column: 20
High quality sequence stop: 545.

FEATURES
source
1..551
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3631603"
/clone_1lb="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(C). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 158 a 129 c 153 g 111 t
ORIGIN

Query Match 12.4%; Score 345; DB 10; Length 551;
Best Local Similarity 99.4%; Pred. No. 2.7e-170;
Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1217 cctgaagtaaacctcattaccacacagagggcgacactggggagagccttattgtg 1276
|||||
Db 34 CCGAAGTCAAAACCTCATTAACCCACAGAGGGCGCACACTGGGAGAGAGCCCTTATGTTG 93
QY 1277 caggagatctggcgctgcttgcagacattacacacctgtagacacagaagagacaca 1336
|||||
Db 94 CAGGGATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTGTAGACACAGAGGACACA 153
QY 1337 ttcaagagagaagccttacctgtcagggagtgtagcaaggttagcagaagtcaca 1396

Db 154 TTGAGAGAGAAACCTTACTTTGACGGAGGTGTGAGCAAGGCTTGACCGAAGTCACA 213
QY 1397 cctcatcagaaccttaagacacacacagagagaagccttattgtatgcagaatgtg 1456
|||||
Db 214 CTTTCATCAGACACTTAAAGGACACACAGAGAGAGCCCTTATGATGACACAGAAATGTG 273
QY 1457 ggcacacttaagcttggaatcaacctcaaaacacacacagagacactcaagggttaa 1516
|||||
Db 274 GCCTCACTTTAGTGTGAATCAAACTCAAAACACACAGAGACACTCAGAGGGTTAA 333
QY 1517 acctatgctgccttgagtgcgggcagltgctttagcctgaagtcacaacttaacaaaca 1576
|||||
Db 334 ACCTTATGCTGCTGGAGTGGCGGAGTGTCTTACCTGAAGTCTAACCTTAAACAACA 393
QY 1577 ccgaaggtcaacacaggggagagaagcattgtatgtagagatgtggcgagccttac 1636
|||||
Db 394 CCGAGAGTCACACCGGGGAGAGCACTTTGTATGTAGGAGTGTGGCGAGCTTTAC 453
QY 1637 ccgaatcaaaccttataccagcaccagagagacacactcagggagagaagcattgtatg 1696
|||||
Db 454 CCGGAATCAACCCCTGAGCAGCAGCAGAGACACACTCAGGGAGAACGCAATTTGTATG 513
QY 1697 tgctgagtgtagagag 1714
|||||
Db 514 TCCTGAGTGTGAGCAGAG 531

RESULT 35
LOCUS A1860707 508 bp mRNA linear EST 07-MAR-2000
DEFINITION w151d12.x1 NCI_CGAP_Ut1 Homo sapiens CDNA clone IMAGE:2424983 3',
mRNA sequence.
ACCESSION A1860707
VERSION A1860707.1 GI:5514323
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
1 (bases 1 to 508)
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskajuk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1940 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 406.

FEATURES
source
1..508
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2424983"
/clone_1lb="NCI_CGAP_Ut1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

BASE COUNT 140 a 110 c 138 g 120 t
ORIGIN

Query Match 12.4%; Score 344; DB 9; Length 508;
Best Local Similarity 99.4%; Pred. No. 9, 1e-170;
Matches 494; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2276 tgaagagaatgcgcgcgcatttcgaagagccgcgcgccttcctcaactgagatgtgag 2335
|||||
DB 508 tgaagagaatgcgcgcgcatttcgaagagccgcgcgccttcctcaactgagatgtgag 449
|||||
QY 2336 ttgctgaaacccgcgcgcgaagtaagtgagcagagagcagtaacatgcgcgcagata 2395
|||||
DB 448 ttgctgaaacccgcgcgcgaagtaagtgagcagagagcagtaacatgcgcgcagata 389
|||||
QY 2396 ggggtggtgactgtgtgaaacccaacctaagcgtgaagacgctccgcgcctaactca 2455
|||||
DB 388 ggggtggtgactgtgtgaaacccaacctaagcgtgaagacgctccgcgcctaactca 329
|||||
QY 2456 tactgaatgaaacgtgcgcgccttcgaatgtgtgtcctccgcgcgaatgcgcgcaccc 2515
|||||
DB 328 tactgaatgaaacgtgcgcgccttcgaatgtgtgtcctccgcgcgaatgcgcgcaccc 269
|||||
QY 2516 tcacctattacgtatgacgtgcgcgccttcctcaatgtgttttcaactgctgtgcaccc 2575
|||||
DB 268 tcacctattacgtatgacgtgcgcgccttcctcaatgtgttttcaactgctgtgcaccc 209
|||||
QY 2576 ttgagtggtgcttgcgtatgacatcaatcagtgacgtgacgtatccctatcttcgagcc 2635
|||||
DB 208 ttgagtggtgcttgcgtatgacatcaatcagtgacgtgacgtatccctatcttcgagcc 149
|||||
QY 2636 cataaagaccagacgtgcgcgcagtgagagagaatcacgcgtgtgagagtgagc 2695
|||||
DB 148 cataaagaccagacgtgcgcgcagtgagagagaatcacgcgtgtgagagtgagc 89
|||||
QY 2696 gacacgtcgc 2755
|||||
DB 88 gacacgtcgc 29
|||||
QY 2756 ctaccatcctcaccc 2772
|||||
DB 28 ctaccatcctcaccc 12
|||||

RESULT 36
AI289399/c 399 bp mRNA linear EST 21-DEC-1998
LOCUS qw32e05.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:1992800 3,
DEFINITION mRNA sequence.
ACCESSION AI289399
VERSION AI289399.1 GI:3931430
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 2932 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 395.

FEATURES
SourceLocation/Qualifiers
1..399

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1992800"
/clone_11b="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-Sport6; Site: 1; Salt;
Site: 2; NCI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 114 a 76 c 113 g 96 t
ORIGIN

Query Match 12.4%; Score 343; DB 9; Length 399;
Best Local Similarity 99.7%; Pred. No. 2, 9e-169;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2375 agtcaaatgcccagcagataggggtgtgtaccgtgtgaacccaacttaagctgaag 2434
|||||
DB 399 agtcaaatgcccagcagataggggtgtgtaccgtgtgaacccaacttaagctgaag 340
|||||
QY 2435 acagtcggcgtcaaatccctacatgaattggaacctgtctccattgtgtgctt 2494
|||||
DB 339 acagtcggcgtcaaatccctacatgaattggaacctgtctccattgtgtgctt 280
|||||
QY 2495 cctccgattgatacccaaccttcacatattacgataccgcgccttcctcaatgtgt 2554
|||||
DB 279 cctccgattgatacccaaccttcacatattacgataccgcgccttcctcaatgtgt 220
|||||
QY 2555 tttaacgt 2614
|||||
DB 219 tttaacgt 160
|||||
QY 2615 tttatccctattctgagcccaataaagaccagacgtgcgcgcgtgagagagaat 2674
|||||
DB 159 tttatccctattctgagcccaataaagaccagacgtgcgcgcgtgagagagaat 100
|||||
QY 2675 caacctgtgtgaggtgtggtggtggtggtggtggtggtggtggtggtggtggt 2734
|||||
DB 99 caacctgtgtgaggtgtggtggtggtggtggtggtggtggtggtggtggtggt 40
|||||

QY 2735 ttgtcccaataaattcttttaccatccca 2768
|||||
DB 39 ttgtcccaataaattcttttaccatccca 6
|||||

RESULT 37
AA218659/c 392 bp mRNA linear EST 12-MAR-1998
LOCUS zq96g04.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:649878 3', mRNA sequence.
ACCESSION AA218659
VERSION AA218659.1 GI:1832767
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

DEFINITION ESTJ311006 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
 AM958936
 ACCESSION AM958936.1 GI:8148620
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 513)
 AUTHORS Hedde, P., Qi, R., Abernathy, K., Dharap, S., Gasparé, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.
 TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 127
 Seq primer: Reverse.
 FEATURES
 source Location/Qualifiers
 1..513
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGE"
 /note="Vector: pBluescriptSkm"
 BASE COUNT 130 a 142 c 102 g 139 t
 ORIGIN

Query Match 12.2%; Score 337; DB 9; Length 513;
 Best Local Similarity 99.7%; Pred. No. 4.5e-166;
 Matches 387; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 2377 tcaaatgcccagcagataggggtggtggtacccgtgtaaacccaacttaagctgaagac 2436
 1 TCAANTGCCAGCAGCATGAGGGGTGGTACCTGCTGAACCCACCTTAAAGCTGAAGAC 60
 QY 2437 agtcccgagtaaatccctcaactgaatgagaacccgtctctccacttggtgtgtctcc 2496
 Db 61 AGTCCCGGCTAAATCTCTACTGATTGAGAACCTGCTTCCATTGGTGGCTTCC 120
 QY 2497 tccgattgaccacccctcaactatttaagtaactgccccttcttaattggtttt 2556
 Db 121 TCCGATTGATCCCAACCTTCACCTATTTTACGTATACCTGCCCTTCTTAATTGGTTT 180
 QY 2557 taacatgctggtcccaactttagtgggtggtccttgcaactaactaactcaactgaactg 2616
 Db 181 TACACTGCTGTGCGCCACCTTTGAGTGTGCTTTCATACCTTACAAATCAGTCAACGTG 240
 QY 2617 tattccctattctgagccataaagaaccagactcaactgaactgagtgagagaataca 2676
 Db 241 TATTCCTCCCTATTCTGAGCCCATAAAGACCCAGACTCAGCTGAGTGAAGAGAGAAATCA 300
 QY 2677 ccttgctggtgaggttggtgggaccactcctgcatccctctcactgaagctgttctt 2736
 Db 301 CCTGCTGTGGGGGTGGGAGCCACTCCCTGCATCCCTCCACATGAGAGCTGTTCTTT 360
 QY 2737 tgcataaataattcttcttaaccatc 2764
 Db 361 TGCTCAATAAATTCCTTCTTACCCATC 388
 RESULT 40
 BI256434 794 bp mRNA linear EST 17-JUL-2001
 LOCUS 602974454F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5113861-5',
 DEFINITION mRNA sequence.
 ACCESSION BI256434
 VERSION BI256434.1 GI:14810838

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 794)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN at: <http://image.llnl.gov>
 Plate: LLAM11276 row: k column: 14
 High quality sequence stop: 791.
 FEATURES
 source Location/Qualifiers
 1..794
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5113861"
 /clone_lib="NIH-MGC_12"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-Sport6; Site-1: NotI; Site-2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."
 BASE COUNT 231 a 209 c 188 g 166 t
 ORIGIN

Query Match 12.0%; Score 333; DB 10; Length 794;
 Best Local Similarity 99.8%; Pred. No. 6.3e-164;
 Matches 453; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Db 333 gaaatccatctcttaacaaaactcaatgctcagctgagcagggagagcgccctg 392
 Db 189 GAAATCCATCTTCTTAACCAAACTCATGCTCAGCTGGAGCGGAGACCGCCCTGG 248
 QY 393 agagaagaagaagaatgctcactgagccctctgccaagatcgagaagcaaatcaact 452
 Db 249 AGGAGGAGAGAAATGCTCACTGGACCTGTGTCCGAATCGAAGCCAGAAATTCACATT 308
 QY 453 agtccctcctgcccctctgatttctccagtcagcaagctctcagccaactgtgtgctg 512
 Db 309 AGTCCCTCCTGCGCTGTGATTTTCTCCAGTCAAGCAAGCTCTCAGCAACATGTGTGCTG 368
 QY 513 agtcatctctcagctgttcttaagttatgggcaagaatcctctccaactgggaaaa 572
 Db 369 AGTCATCTCTCAGCTGTTTTCAGTTATGGGCGAAGAACTCTCCACCTGGGAAAA 428
 QY 573 cactatcagaagaataagaacaacagcagagatccattctgctttagtggcaagaaga 632
 Db 429 CACTATTCAGAAATCAGAAACACAGACAGATTCATTGCTTGTGGAAGAGAGAA 488
 QY 633 tggattcaagaagga-gaagactccagactcctgttctggagagtaagcaaaatgagac 691
 Db 489 TGGATTCAAGAGGAGGAGAGAGACTCCAGACTCCGTGTTGGGAGAGTAAAGAAATGGCAC 548
 QY 692 ttcaaaagcaattccagccacactgaaagaacaacagccaagcagatccaaggaagaca 751
 Db 549 TTCAAAGGCACTTCCAGGCCACCTGAAGAACACAGCAGCAGCACTCCAAAGAGACAA 608
 QY 752 cacagctggtgatataggttccagccctgaagc 785
 Db 609 CACAGTGTGTGATATAGGGTCCAGCCCTGACCG 642

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	ORIGIN
41	AA450118/c	383 bp	mRNA	linear	EST 04-JUN-1997									
	AA450118	2x42609.s1 Soares_total_fetus_Nb2HF8_9w	Homo sapiens	CDNA clone										
	IMAGE:789160 3'	mRNA sequence.												
	AA450118	AA450118.1	GI:2163868	EST.										
	ORGANISM	human.												
	SOURCE	Homo sapiens												
	REFERENCE	Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Euteleostomi;												
	AUTHORS	1 (bases 1 to 383)												
		Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,												
		Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,												
		Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,												
		, T., Waterston, R., and Wilson, R.												
		Washu-Merck EST Project 1997												
		Unpublished (1997)												
		Contact: Wilson RK												
		Washington University School of Medicine												
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108												
		Tel: 314 286 1800												
		Fax: 314 286 1810												
		Email: est@watson.wustl.edu												
		This clone is available royalty-free through LNL ; contact the												
		IMAGE Consortium (info@image.llnl.gov) for further information.												
		Seq primer: -41ml3 fwd. ET from Amersham												
		High quality sequence stop: 352.												
		Location/Qualifiers												
		1..383												
		/organism="Homo sapiens"												
		/db_xref="GDB:5986777"												
		/db_xref="taxon:9606"												
		/clone="IMAGE:789160"												
		/clone_lib="Soares_total_fetus_Nb2HF8_9w"												
		/dev_stage="8-9 weeks"												
		/lab_host="DH10B"												
		/note="Vector: pT730-Pac (Pharmacia) with a modified												
		polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand												
		was prepared from mRNA obtained from pooled 8-9 week												
		(total) fetus material with a Not I - oligo(dT) primer [5'												
		TGTTACCAATCTGAAGTGGGAGCGCCGCTTAATTTTTTTTTTTT 3'].												
		Double-stranded cDNA was ligated to Eco RI adaptors												
		(Pharmacia), digested with Not I and cloned into the Not I												
		and Eco RI sites of the modified pT73 vector. Library												

	Db	143	TATTCGAGCCCATTAAGAACCACCACTGACTGCAGTGAAGAGAAATCACCTGCTG	84
Oy	2685	tggaggttggggaacacatcccgcatccttcaccattgcacttgagagctgtctttgtccaat	2744	
Db	83	TGGGGGTGGGGAGCACCACCTCCCGCATCCCTTCACAGTAGAGCTGTCTTTGGCTCAAT	24	
Oy	2745	aaaatctttctaccacctc 2767		
Db	23	AAAATTCTTTCTACCCATCTTC 1		
RESULT 42				
LOCUS	AA84745/c	378 bp	mRNA	linear EST 19-AUG-1997
DEFINITION	neg1h06.s1 NCI_CGAP_Ew1 Homo sapiens CDNA clone IMAGE:910715, mRNA sequence.			
ACCESSION	AA84745			
VERSION	AA84745.1	GI:2214130		
KEYWORDS	EST.			
SOURCE ORGANISM	Homo sapiens human. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE AUTHORS TITLE	1 (bases 1 to 378) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)			
JOURNAL COMMENT	Contact: Robert Strusberg, Ph.D. Email: cgabs@emall.nih.gov Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CNLA Library Preparation: David B. Krizman, Ph.D. CNLA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bdrip/image/image.html Insert Length: 457 Std Error: 0.00 Seq primer: -41m13 fwd. RT from Amersham High quality sequence stop: 377. Location/Qualifies			
FEATURES source	1..378 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:910715" /clone_lib="NCI_CGAP_Ew1" /tissue_type="Ewing's sarcoma" /lab_host="DH10B" /note="Vector: pAMP10; mRNA made from Ewing's sarcoma, CDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."			
BASE COUNT	111 a 67 c 108 g 92 t			
ORIGIN				
Query Match	11.5%;	Score 318;	DB 9;	Length 378;
Best Local Similarity	99.7%;	Pred. NO. 4.5e-15e;		
Matches 368;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Oy	2400	tgggtaacctggygaaccaacttaagaagcygaagaacagtcccggctaataatcataat	2459	
Db	378	TGGGTAACCTGGTGGAACCAACCACTTAAGAGCTGAAGACAGTCCCGCTAATTCATACT	319	
Oy	2460	gaattgagaacctgtcttctcccaatttgtgtgtcttctccgatltgatgccaacacctcac	2519	
Db	318	GAATTGAGACAACCTGCTTCCATTTGGTGGTGTCTTCTCCGATGATGCCAACACCTTCAC	259	
Oy	2520	ctattttacgtataacctgtccctttcccaatttggtttttaaacatgtctgtgccacttttt	2579	

Dd	258	CTATTTCACGATACACTGCCTTTTCCATTAATGGTTTTAAACACTGTGTGGCCACCCTTTTG	199
Oy	2580	agtgtgccttgactgaacttaacaatcagtcacgtagtattccctattctcgagcca	2639
Dd	198	AAGTGTGCCCTTGGCACTTTCMAAATCAGTCACACTGTATTCCTCATTTCTGAGCCCCATA	139
Oy	2640	aagaagcccaagactgagctgcagtagagagaanaatcacccctgcctctggagttggagacc	2699
Dd	138	AAAGACCCAGACTCAGCTGCAGTAGAGAGAGAATACCCCTGCTGTGGGGGTGGGACC	79
Oy	2700	acctccctgatccccctcccacactgagagactgtctctcttgcataaaatctttcttac	2759
Dd	78	ACTCCCTCGATGCCCTTCACAAGTAGACTGTCTTTTGGCTCAATAAATCTTTTCTAC	19
Oy	2760	cocatcccta	2768
Dd	18	CCATCTCA	10
RESULT	43		
LOCUS	A1159662/c		
DEFINITION	A1159662	369 bp	mRNA linear EST 26-Oct-1998
ACCESSION	gb08f10.x1 Soares_fetal_heart_NbHH19W	Homo sapiens	cdna clone
KEYWORDS	A1159962		
SOURCE	A1159962.1	GI:3693342	
ORGANISM	EST.		
JOURNAL	human.		
COMMENT	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCI-Cgap http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov This clone is available royalty-free through LINT ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 655 Std Error: 0.00 Seq primer: -40m13 fwd. RT from Amersham High quality sequence stop: 364. Location/Qualifiers 1. 369 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="1706443" /clone_lib="Soares_fetal_heart_NbHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" /note="Organ: heart; Vector: pTZ19D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' TGTATCACATCTGAAGTGGAGCGGCCCATCTTTTCTTTTCTTTTCTT 3'] double-stranded CDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pTZ19 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Felina Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHH19W."		
BASE COUNT	109 a	70 c	105 g 85 t
ORIGIN			
Query Match	11.4%; Score 315; DB 9; Length 369;		
Best Local Similarity	99.7%; Pred. No. 1.7e-154;		
Matches	365; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Oy	2397	gggtgggtactctgtgtgaaccaacttaagctgaagacagtcgccgctaattcccat	2456

Db	369	GGGTGGTACTGTGGGAACCCCAACCTTAAAGCTGAAACAGTCCGGCTTAATCTCAT	310
QY	2457	actgaattgagaacactgctctccatcttggtgctcttcctccgatgatgatacccaacct	2516
Db	309	ACTGAAATGAAGAACCTGCTCTCCATTTGGTGCTCTTCCATGATGATCCCAACCTT	250
QY	2517	caacctatttgcgtatatactgctccttccctaattggtttttaacctgctgtgccacct	2576
Db	249	CACCAATTTTACGATATACCTCCCTTCCATTTGATTTTACACTGCTGTGCCACCTT	190
QY	2577	ttgagtgtgctcttgatctgatacttaacaatgaactgaactgfatccctattccctgagccc	2636
Db	189	TTGATGTGCTCTTTGCACTTACAAATCAGTCACTGATTTCCCTATTTCTGAGCC	130
QY	2637	ataaagaagccagactcagctcagctgagtgagagagaataacacctgctgtgtaggtggg	2696
Db	129	ATMAAGAGCCAGACTAGCTGAGTGAAGAGAGAAATCACCTGCTGTGGGGTGGG	70
QY	2697	accactccctgataccctccctccactgagagctgctcttctgtcctaataatctcttc	2756
Db	69	ACCACCTCCCTGATCCCTCTCCACTGAGACCTGTTCTTTGCTCAATAAATCTTTTC	10
QY	2757	taccga 2762	
Db	9	TACCA 4	
RESULT	44		
AL566746/C			
LOCUS	AL566746	818 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL566746 LTI_FL013_FBrn1	Human sapiens	CDNA clone CS0DF025Y022 3
ACCESSION	AL566746		
VERSION	AL566746.1	GI:12919412	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
FEATURES	Genoscope - Centre National de Sequencage		
source	BP 191 91006 EVRY cedex - France		
	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
	1. 818		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DF025Y022"		
	/clone_lib="LTI_FL013_FBrn1"		
	/dev_stage="pooled tissue from post conception fetuses (20		
	week, 24 week and 26 week)"		
	/lab_host="DH10B"		
	/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand		
	CDNA was primed with a NotI-oligo(dT) primer. Five prime		
	end enriched, double-stranded cDNA was digested with Not I		
	and cloned into the Not I and Eco RV sites of the		
	pCMVSPORT 6 vector. Library was constructed by Life		
	Technologies. Contact : Feng Liang Life Technologies, a		
	division of Invitrogen 9800 Medical Center Drive Rockville		
	, Maryland 20850, USA Fax : (1) 301 610 8371 Email :		
	fliang@lifestech.com URL :		
	http://fulllength.invitrogen.com"		
BASE COUNT	199 a	208 c	190 g
ORIGIN			212 t
			9 others
Query Match	11.4%	Score 315,	DB 9, Length 818;
Best Local Similarity	99.2%;	Pred. No. 2e-154;	

Matches 635; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1855 tctgcagagggatgtggaagagccttctgtcgaatgaactcgaatgaacaccagagag 1914
|||||
Db 692 tctgcagagggatgtggaagagccttctgtcgaatgaactcgaatgaacaccagagag 633
QY 1915 cacacgaggggggaagcctcattgtgtcagagagagtgaggcagagccttaagcggcagt 1974
|||||
Db 632 cacacgaggggggaagcctcattgtgtcagagagagtgaggcagagccttaagcggcagt 573
QY 1975 caacactcattgaacaccagagacacattcagagagagagccttatttgcagaagt 2034
|||||
Db 572 caacactcattgaacaccagagacacattcagagagagagccttatttgcagaagt 513
QY 2035 gtgcagagggccttaagtcgaagtc--caacattcagacatagaagagacacactcagga 2093
|||||
Db 512 gtgcagagggcctttagtcgaagtc--caacattcagacatagaagagacacactcagga 453
QY 2094 tagaacttattgtatagaggaattgtgtacagaccttagccagagagtcatactcattc 2153
|||||
Db 452 tagaacttattgtatagaggaattgtgtacagaccttagccagagagtcatactcattc 393
QY 2154 agacaccagagagacacacacagtcgtgtgtgttccagccattgtatagaccagaagt 2213
|||||
Db 392 agacaccagagagacacacacagtcgtgtgtgttccagccattgtatagaccagaagt 333
QY 2214 gagaactcgtgtgtgtatgacatgagagacgttactgtgaagactgtatctcaccaga 2273
|||||
Db 332 gagaactcgtgtgtgtatgacatgagagacgttactgtgaagactgtatctcaccaga 273
QY 2274 cctgaagaggaattgtgtgtcatttccaggaagcctgtccctcctcactgtgtgtgtgt 2333
|||||
Db 272 cctgaagaggaattgtgtgtcatttccaggaagcctgtccctcctcactgtgtgtgtgt 213
QY 2334 ggtgtgtgaagaccggttcaggtatgtagtgcagagagagcagtcgaatgtccagagcaga 2393
|||||
Db 212 ggtgtgtgaagaccggttcaggtatgtagtgcagagagagcagtcgaatgtccagagcaga 153
QY 2394 taagggt 2453
|||||
Db 152 taagggt 93
QY 2454 cactgtgaattgagaacctgtcttccattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2493
|||||
Db 92 cactgtgaattgagaacctgtcttccattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 53

RESULT 45
R50617 501 bp mRNA linear EST 18-MAY-1995
LOCUS yj57b01.r1 Soares breast 2ndHst Homo sapiens cDNA clone
DEFINITION IMAGE:152809 5' similar to SP:KRL_HUMAN P10072 HKRI PROTEIN ;, mRNA
sequence.
ACCESSION R50617
VERSION R50617.1 GI:812519
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marrs,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
JOURNAL
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelw@wustl.edu

Insert Size: 770
High quality sequence stops: 323 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 770 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 323.
Location/Qualifiers
1. 501
/organism="Homo sapiens"
/db_xref="GDB:565059"
/db_xref="taxon:9606"
/clone="IMAGE:152809"
/clone_1b="Soares breast 2ndHst"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCAGATGAGGAGCGGCGCCCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT73 vector (Pharmacia).
Library went through one round of normalization to a Col -
230. Library constructed by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 115 a 113 c 118 g 149 t 6 others
ORIGIN

Query Match 11.3%; Score 312; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 6.9e-153;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2078 gaggacacactcagagatagaacttattgtatagagagtggtgtacagccttagcca 2137
|||||
Db 1 GAGGACACACTCAGATGAGAACTTATGTATAGGAATGTGTACAGCTTTAGCCA 60
QY 2138 ggaatcattctcattcagacacagagacacacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2197
|||||
Db 61 GGAGTCATCTTCATCAGACCCAGAGACACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 2198 gctagataccagagtgagacattctgtgtgtatgatcatgatgagactgtactgttaaga 2257
|||||
Db 121 GCTAGATACCAAGTGGAGACATTCGCTGTGATATGATGAGAGCTGTACTGTGTAAGA 180
QY 2258 ctgtatctccatccacactgaaggaattgtgtgtcatttccaggaagcctgtcccttc 2317
|||||
Db 181 CTTGTATCTCCATCCACCTGGAAGGAATTCGTGCTCATTTTCAGAGACCTGCTTC 240
QY 2318 ctccactgt 2377
|||||
Db 241 CTGACTGT 300
QY 2378 caaatgtccagg 2389
|||||
Db 301 CAATGCCAGG 312

RESULT 46
R83624 427 bp mRNA linear EST 04-AUG-1995
LOCUS ypl6h01.r1 Soares breast 3rdHst Homo sapiens cDNA clone
DEFINITION IMAGE:187633 5' similar to SP:KRL_HUMAN P10072 HKRI PROTEIN ;, mRNA
sequence.
ACCESSION R83624
VERSION R83624.1 GI:928501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Db 110 TCAGNTCAGTAGAGAGAGAAATCCCTGCTGTGGGGGTTGGGACCACCTCCCTGCATC 51
 Oy 2712 cccctccactgagagctgtcttcttgcataaataattcttacc 2761
 Db 50 CCCTCTCCACTGAGAGCTGTTCTTGTCTCAATAAATCTTTCTACCC 1

RESULT 48
 AM139497/c
 LOCUS 383 bp mRNA linear EST 30-OCT-1999
 DEFINITION UI-H-BII-a4e-e-08-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
 IMAGE:2716526 3', mRNA sequence.
 ACCESSION AM139497
 VERSION AM139497.1 GI:6144215
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 383)
 NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
 NCI-CGAP clone distribution information can be found through the
 I.M.A.G.E. Consortium/ILNI at:
 www.bio.illn.gov/dbrr/image/image.html The following repetitive
 elements were found in this cDNA sequence: 52-178, >MER524unknown
 Seq primer: M13 Forward
 POLYA-Tes.

FEATURES
 source
 Location/Qualifiers
 1..383
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2716526"
 /clone_lib="NCI_CGAP_Sub3"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NCI_CGAP_Sub3 library is a subtracted library derived from
 the NCI_CGAP_Sub1 library, which is a subtracted library
 derived from B1. B1 constitutes a mixture of 21
 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_C04
 , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_C010,
 NCI_CGAP_C016, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
 NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
 NCI_CGAP_Br2, NCI_CGAP_C08, NCI_CGAP_C111, NCI_CGAP_Lym2,
 NCI_CGAP_Br23, NCI_CGAP_C08, NCI_CGAP_C111, NCI_CGAP_Lym2,
 NCI_CGAP_Br23, NCI_CGAP_Ly5, NCI_CGAP_Ly44,
 NCI_CGAP_Br25, NCI_CGAP_G04, NCI_CGAP_G06,
 NCI_CGAP_Br25. These 21 libraries were pooled and a
 single-stranded DNA preparation of the resulting mixture
 was used as a tracer in a subtractive hybridization with
 a driver whose composition is detailed below:
 NCI_CGAP_Kid3 pool 1 LHAM 3334-3337, 3682-3683,
 3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775
 , 1500552-1502855); NCI_CGAP_Kid5 pool 1 LHAM 3338-3342
 , 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831,
 1471368-1472803, 1492104-1493255); NCI_CGAP_Ly5 pool 1
 LHAM 3575-3582, 3851-3854 (IMAGE Clonoids 141920-1417991,
 1520904-1522439); NCI_CGAP_G04 pool 1 LHAM 3164-3167,
 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,
 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
 LHAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids
 985608-986758, 1101192-1101959, 1217928-1220615);
 NCI_CGAP_C010 pool 1 LHAM 2644-2653, 2871-2872 (IMAGE
 Clonoids 1057416-1061255, 1144584-1145351). Subtraction
 was performed as previously described [Bonaldo, Lennon &

BASE COUNT 107 a 69 c 107 g 100 t
 ORIGIN

Query Match 11.0%; Score 305; DB 9; Length 383;
 Approaches To Facilitate Gene Discovery. Genome Research
 6: 791-806.
 Best Local Similarity 99.7%; Pred. No. 3.2e-149;
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 2409 gctgaaccccaacttaaaactgaagacagtcgccgagctaaactcactgaattaga 2468
 Db 373 GGTGAACCCACACTTAACACTGAAGACACATCCCGGTAATCTCTACTGATTTAGA 314
 Y 2469 acctgtctccatttgltgtcttccctccgattgatcccaacctcactatttac 2528
 Db 313 ACCGTCTTCCATTGGTGTGCTTCCCGATGATGCCAACCCCTTCACTATTTTAC 254
 Y 2529 gtatactgcccttccctaattggttttacaactgtctgtccaccttttgagtgtgc 2588
 Db 253 GTATACCTGCCCTTCTTAATTGGTTTACACTGCTGCCACCTTTGAGTGTGCC 194
 Y 2589 ttgcactactaacaatcagtaacagtgatctccctatctgagcccaaaagacca 2648
 Db 193 TTTGCACTACTTACAAACAGTCAACGATGATCCCTATTCTGAGCCATTAAGACCA 134
 Y 2649 gactcagctgcaatgagagagaataaccctctgtgagagttgggagcaactccctgc 2708
 Db 133 GACTCAGCTCAGAGAGAGAGAAATCACCCTGCTGGGGGTTGGGAGACCACTCCCTGC 74
 Y 2709 atcccctccactgagagctgtcttcttgcataaataattcttaccatc 2764
 Db 73 ATCCCTCTCCACTGAGAGCTGTTCTTGTCTCAATAAATCTTTCTACCATC 18

RESULT 49
 AM579734
 LOCUS 680 bp mRNA linear EST 16-MAR-2000
 DEFINITION RCI-HT0256-080100-012-a08 HT0256 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM579734
 VERSION AM579734.1 GI:7254783
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 680)
 HCGP http://www.judwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@judwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.judwig.org.br/scripts/gethtml2.pl?l=RCI&t2=RCI-HT0256-
 080100-012-a08&t3=2000-01-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 22
 High quality sequence stop: 661.

FEATURES
 source
 Location/Qualifiers
 1..680
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone.lib="Hf0256"
/dev_stage="Adult"
/note="Organ: head,neck; Vector: puc18; Site_1: Sma1;
Site_2: Sma1: A mini-library was made by cloning products
derived from ORESFES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 205 a 140 c 159 g 176 t
ORIGIN

Query Match 10.5%; Score 291; DB 9; Length 680;
Best Local Similarity 99.7%; Pred. No. 8.7e-142;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1990 accagagagacattcagagagagacattatcttcagagagtgtagcgagcttta 2049
|||||
DB 291 ACCAGAGACACATTCAGAGAGAGACGCTTATTTGCAAGAGTGTGAGCGGCTTTA 350
QY 2050 gtgcgagatccacattatcagacatcagagagacacattagagatagaacttatgtt 2109
|||||
DB 351 GTGCGAATGTGTTACAGCTTATCAGACATCAGAGACACTGAGATGAAACTTTATGTGT 410
QY 2110 atagggaaatgtgttacagaccttagcagagagatcattatcagacacagagagac 2169
|||||
DB 411 ATAGGGAAATGTGTTACAGCTTATCAGAGAGATCATTACTTATCAGACACAGAGACAC 470
QY 2170 acacagatgtctgtgtcttttcagacattgtctagatacacaagtgtgagacattgtgtgt 2229
|||||
DB 471 ACCAGATGCTGTGCTTTTACACCATGCTAGATACCAAGTGAGACATCTGTGTGT 530
QY 2230 gattatgcatgagacgtactgttgaagactgtatctccaccactgaaggagaaattgc 2289
|||||
DB 531 GATTATGACATGACATGTTAGTGTGAAGACTGTATCTCCATCCAGTGAAGGAAATGTC 590
QY 2290 tggctcatttcagagagccctgccttcctcactgtgagatg 2331
|||||
DB 591 TGGCTCATTTTTCAGGAGCCCTGCCCTTCCCTCACTGTGGATGG 632

RESULT 50
BE798936 1041 bp mRNA linear EST 20-SEP-2000
LOCUS 601583719F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3937924 5',
DEFINITION mRNA sequence.
ACCESSION BE798936
VERSION BE798936.1 GI:10220134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 1041)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LLCM784 row: f column: 05
High quality sequence start: 58
High quality sequence stop: 705.
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:3937924"
/clone.lib="NIH_MGC_7"
/tissue_type="Small cell carcinoma"
/cell_line="MG3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dr priming. Directionally
cloned into ECORI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 320 a 276 c 281 g 164 t
ORIGIN

Query Match 10.5%; Score 290; DB 10; Length 1041;
Best Local Similarity 100.0%; Pred. No. 3.2e-141;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1364 gtagtgtagcaaggttttagccagaaatcacaacctcagacacttaagagacacac 1423
|||||
DB 113 GGAGTGTGAGCAAGGCTTTAGCCAGAGTCACACTCATCAGACACTTAAGACACACAC 172
QY 1424 agagagaaagccttatgtatgacagaaatgtggcgtcacttagctggaatcaaacct 1483
|||||
DB 173 AGGAGGAAGACCTTATGTATGACACAGATGTGGCGCTTATAGCTGGAATCAAACT 232
QY 1484 caaaacacaccagagagacacactcaggggttaaaccttatgtctgctgtgagtgagca 1543
|||||
DB 233 CAAACACACACAGAGAGACACACTCAGGGGCTTAAACCTTATGTCTGCTGGAGTGGCGGCA 292
QY 1544 gtgccttagccttaaatcaaaccttaacaaacccagaggtcacacagggggagaaagcc 1603
|||||
DB 293 GTGCTTTAGCTTAAGTCAAACCTTAAACACACAGAGGTCAACACGAGGAGAAAGCC 352
QY 1604 attgtatgtacgagatgtgtggcagagccttacccggaatcaaacctga 1653
|||||
DB 353 ATTGTATGTAGCGAGTGTGGGAGGCTTTACCGGAATCAACCTGA 402

Search completed: May 16, 2002, 09:13:25
Job time: 3170 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:20:45 ; Search time 4965.75 seconds
(without alignments)
11681.691 Million cell updates/sec

Title: US-09-898-556A-3
2772
Perfect score: 1 cagcgcgcttaagctggtg.....ttctaccaccctcaccct 2772
Sequence:

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Match	Length	ID	Description
--------	-------	-------	--------	----	-------------

1	2772	100.0	2772	9	AB013897	AB013897 Homo sapi
2	2383	86.0	2854	9	AK056243	AK056243 Homo sapi
3	2244	81.0	135173	9	AC008806	AC008806 Homo sapi
4	2244	81.0	253217	2	AC016590	AC016590 Homo sapi
5	2193	79.1	38173	9	AC093235	AC093235 Homo sapi
6	1771	63.9	1998	9	BC004513	BC004513 Homo sapi
7	672	24.2	672	9	HUMHMR1	M20675 Human Krupp
8	571	20.6	678	6	AX067361	AX067361 Sequence
9	37	1.3	130351	2	AC024911	AC024911 Homo sapi
10	37	1.3	161625	2	AC096898	AC096898 Homo sapi
11	37	1.3	164264	2	AC034197	AC034197 Homo sapi
12	37	1.3	165379	2	AC069265	AC069265 Homo sapi
13	37	1.3	184375	2	AC026219	AC026219 Homo sapi
14	36	1.3	148270	2	AC027579	AC027579 Homo sapi
15	36	1.3	174405	2	AC022727	AC022727 Homo sapi
16	36	1.3	179792	2	AP001531	AP001531 Homo sapi
17	34	1.2	150794	2	AC025211	AC025211 Homo sapi
18	34	1.2	166650	2	AC087493	AC087493 Homo sapi
19	34	1.2	218922	2	AC008106	AC008106 Homo sapi
20	32	1.2	162453	2	AC027607	AC027607 Homo sapi
21	32	1.2	171279	2	AC036185	AC036185 Homo sapi
22	32	1.2	171601	2	AC026269	AC026269 Homo sapi
23	32	1.2	173133	2	AC016166	AC016166 Homo sapi
24	32	1.2	174701	2	AC098581	AC098581 Homo sapi
25	32	1.2	175833	2	AC073412	AC073412 Homo sapi
26	32	1.2	207362	2	AC007380	AC007380 Homo sapi
27	31	1.1	154746	9	AL157944	AL157944 Human DNA
28	31	1.1	168736	2	AC019042	AC019042 Homo sapi
29	31	1.1	177864	9	AC024996	AC024996 Homo sapi
30	29	1.0	2328	9	HSW800283	AL049942 Homo sapi
31	29	1.0	2790	9	AK056466	AK056466 Homo sapi
32	29	1.0	3178	9	AK056184	AK056184 Homo sapi
33	29	1.0	3273	9	BC021298	BC021298 Homo sapi
34	29	1.0	127867	2	HS694814	AL031673 Human DNA
35	29	1.0	170129	2	AC109473	AC109473 Homo sapi
36	28	1.0	181770	2	BC011862	BC011862 Homo sapi
37	28	1.0	2823	9	AK055734	AK055734 Homo sapi
38	28	1.0	3055	9	AC091936	AC091936 Homo sapi
39	28	1.0	61452	2	AF161800	AF161800 Homo sapi
40	28	1.0	62002	2	AC105030	AC105030 Homo sapi
41	28	1.0	65236	2	AL512489	AL512489 Human DNA
42	28	1.0	80846	9	HSJ734P14	AL049650 Human DNA
43	28	1.0	90220	9	AF159227	AF159227 Homo sapi
44	28	1.0	109217	9	AC099798	AC099798 Homo sapi
45	28	1.0	131328	9	AC011059	AC011059 Homo sapi
46	28	1.0	143799	9	AC010998	AC010998 Homo sapi
47	28	1.0	146831	2	AC021765	AC021765 Homo sapi
48	28	1.0	152393	2	AC104974	AC104974 Homo sapi
49	28	1.0	154604	9	AC080121	AC080121 Homo sapi
50	28	1.0	154987	9	AC092333	AC092333 Homo sapi
51	28	1.0	157402	9	AC012642	AC012642 Homo sapi
52	28	1.0	157478	2	AC015515	AC015515 Homo sapi
53	28	1.0	158039	2	AC021840	AC021840 Homo sapi
54	28	1.0	158113	2	AC011159	AC011159 Homo sapi
55	28	1.0	158999	2	AC083843	AC083843 Homo sapi
56	28	1.0	159698	9	AC027269	AC027269 Homo sapi
57	28	1.0	159956	9	AC012642	AC012642 Homo sapi
58	28	1.0	161274	9	AC012690	AC012690 Homo sapi
59	28	1.0	162671	2	AC034299	AC034299 Homo sapi
60	28	1.0	164179	9	AC012060	AC012060 Homo sapi
61	28	1.0	167073	9	AC018465	AC018465 Homo sapi
62	28	1.0	191370	30	AC017018	AC017018 Homo sapi
63	28	1.0	194236	2	AL645944	AL645944 Homo sapi
64	28	1.0	203563	2	AC069454	AC069454 Homo sapi
65	28	1.0	287571	2	AC092338	AC092338 Homo sapi
66	28	1.0	297201	2	AC091956	AC091956 Homo sapi
67	27	1.0	1667	9	HSZNF169S2	U28251 Human Krupp
68	27	1.0	1945	9	BC008213	BC008213 Homo sapi
69	27	1.0	3099	9	AF114817	AF114817 Homo sapi
70	27	1.0	3373	9	BC006247	BC006247 Homo sapi
71	27	1.0	3377	9	BC005859	BC005859 Homo sapi
72	27	1.0	25560	9	AP000298	AP000298 Homo sapi
73	27	1.0	40791	9	AF039907	AF039907 Homo sapi

74	27	1.0	100000	9	AP000044	AP000044 Homo sapi	147	25	0.9	182168	2	AL133542	AL133542 Homo sapi
75	27	1.0	100000	9	AP000112	AP000112 Homo sapi	148	25	0.9	186839	2	CNS01DXD	AL133227 Homo sapi
76	27	1.0	100000	9	AP000188	AP000188 Homo sapi	C 149	25	0.9	190005	9	AC068061	AC068061 Homo sapi
77	27	1.0	105779	9	AL137007	AL137007 Human DNA	150	25	0.9	194055	2	AC109480	AC109480 Homo sapi
78	27	1.0	137848	9	AC097479	AC097479 Homo sapi	C 151	25	0.9	197992	2	AC1020715	AC1020715 Homo sapi
79	27	1.0	141843	9	AC105267	AC105267 Homo sapi	152	25	0.9	210107	9	AL365214	AL365214 Homo sapi
80	27	1.0	145294	9	AC009464	AC009464 Homo sapi	153	25	0.9	340000	9	AP001693	AP001693 Homo sapi
81	27	1.0	151458	30	AC021217	AC021217 Homo sapi	C 154	25	0.9	340000	9	AP001715	AP001715 Homo sapi
82	27	1.0	153619	9	AC026239	AC026239 Homo sapi	C 155	24	0.9	2021	5	AY032997	AY032997 Homo sapi
83	27	1.0	153851	2	AC107419	AC107419 Homo sapi	156	24	0.9	2284	9	ALU57453	ALU57453 Xenoopus lae
84	27	1.0	155976	2	AC055113	AC055113 Homo sapi	157	24	0.9	25816	9	AF275816	AF275816 Homo sapi
85	27	1.0	172177	2	AC104190	AC104190 Homo sapi	C 158	24	0.9	26344	9	AP000202	AP000202 Homo sapi
86	27	1.0	183928	2	AL160055	AL160055 Homo sapi	159	24	0.9	58384	9	AP000241	AP000241 Homo sapi
87	27	1.0	184664	9	AC096757	AC096757 Homo sapi	160	24	0.9	67870	2	AC083787	AC083787 Homo sapi
88	27	1.0	222794	9	AC093655	AC093655 Homo sapi	161	24	0.9	71067	2	AF270510	AF270510 Homo sapi
89	27	1.0	340000	9	AP001716	AP001716 Homo sapi	C 162	24	0.9	83412	2	AC106778	AC106778 Homo sapi
90	26	0.9	50343	2	AC091190	AC091190 Homo sapi	C 163	24	0.9	85151	9	AL596219	AL596219 Homo sapi
91	26	0.9	66748	2	AC090817	AC090817 Homo sapi	C 164	24	0.9	85766	9	AL391559	AL391559 Homo sapi
92	26	0.9	11927	2	AC005062	AC005062 Homo sapi	C 165	24	0.9	86156	2	AL162262	AL162262 Homo sapi
93	26	0.9	145495	2	CNS01DS4	AC093273 Homo sapi	C 166	24	0.9	86371	9	AL355672	AL355672 Homo sapi
94	26	0.9	153624	2	CNS01DS4	AL121654 BAC seque	C 167	24	0.9	99257	9	AL355230	AL355230 Human DNA
95	26	0.9	154199	2	AC013639	AC013639 Homo sapi	C 168	24	0.9	101618	9	AC068863	AC068863 Homo sapi
96	26	0.9	165590	2	AL512885	AL512885 Human DNA	C 169	24	0.9	107033	2	AC096949	AC096949 Homo sapi
97	26	0.9	166192	9	AC096737	AC096737 Homo sapi	C 170	24	0.9	107898	9	AF124731	AF124731 Homo sapi
98	26	0.9	169402	2	AC027715	AC027715 Homo sapi	C 171	24	0.9	126344	9	AP000097	AP000097 Homo sapi
99	26	0.9	170378	2	AC013267	AC013267 Homo sapi	C 172	24	0.9	126590	2	AP001112	AP001112 Homo sapi
100	26	0.9	171535	9	CNS01DR2	AL121595 Human chr	C 173	24	0.9	136371	9	AC025451	AC025451 Homo sapi
101	26	0.9	176872	2	AC022302	AC022302 Homo sapi	C 174	24	0.9	138224	9	AC095562	AC095562 Homo sapi
102	26	0.9	178348	9	AP000501	AP000501 Homo sapi	C 175	24	0.9	139930	2	AL160009	AL160009 Homo sapi
103	26	0.9	185209	2	AL390314	AL390314 Homo sapi	C 176	24	0.9	141488	9	AC010256	AC010256 Homo sapi
104	26	0.9	186271	2	AC024168	AC024168 Homo sapi	C 177	24	0.9	146469	2	AC068330	AC068330 Homo sapi
105	26	0.9	186735	2	AC093207	AC093207 Homo sapi	C 178	24	0.9	148697	9	AC010432	AC010432 Homo sapi
106	26	0.9	188921	2	AC025122	AC025122 Homo sapi	C 179	24	0.9	160042	2	AC027757	AC027757 Homo sapi
107	26	0.9	188921	2	AC019339	AC019339 Homo sapi	C 180	24	0.9	160274	2	AC073411	AC073411 Homo sapi
108	26	0.9	198084	2	AC093117	AC093117 Homo sapi	C 181	24	0.9	160457	9	AC006016	AC006016 Homo sapi
109	26	0.9	200426	2	AC012205	AC012205 Homo sapi	C 182	24	0.9	161222	2	AC016037	AC016037 Homo sapi
110	26	0.9	228434	2	AC024898	AC024898 Homo sapi	C 183	24	0.9	162805	2	AC020565	AC020565 Homo sapi
111	26	0.9	243887	2	AX165270	AX165270 Sequence	C 184	24	0.9	163891	9	AC092647	AC092647 Homo sapi
112	25	0.9	51	6	AX165674	AX165674 Sequence	C 185	24	0.9	165616	9	AC073333	AC073333 Homo sapi
113	25	0.9	51	6	AK027183	AK027183 Homo sapi	C 186	24	0.9	166116	2	AC067963	AC067963 Homo sapi
114	25	0.9	3104	9	HS10L11	AL109616 Homo sapi	C 187	24	0.9	169554	2	AC021379	AC021379 Homo sapi
115	25	0.9	40018	9	HS10L11	AL109616 Homo sapi	C 188	24	0.9	169586	9	AC078883	AC078883 Homo sapi
116	25	0.9	58445	9	AL356735	AL356735 Human DNA	189	24	0.9	170847	2	AF188028	AF188028 Homo sapi
117	25	0.9	59951	9	AL356735	AL356735 Human DNA	C 190	24	0.9	172613	9	AC087667	AC087667 Homo sapi
118	25	0.9	66129	9	AP000281	AP000281 Homo sapi	C 191	24	0.9	172759	2	AP0068975	AP0068975 Homo sapi
119	25	0.9	89323	9	AP000021	AL512789 Homo sapi	C 192	24	0.9	175697	2	AP002512	AP002512 Homo sapi
120	25	0.9	97433	2	AL512789	AL512789 Homo sapi	C 193	24	0.9	178064	2	AC023311	AC023311 Homo sapi
121	25	0.9	100000	9	AP000039	AP000039 Homo sapi	194	24	0.9	179144	2	AL606469	AL606469 Homo sapi
122	25	0.9	100000	9	AP000085	AP000085 Homo sapi	C 195	24	0.9	181065	2	AC104989	AC104989 Homo sapi
123	25	0.9	100000	9	AP000107	AP000107 Homo sapi	C 196	24	0.9	181731	9	AC021678	AC021678 Homo sapi
124	25	0.9	100000	9	AP000137	AP000137 Homo sapi	197	24	0.9	183965	9	AC068764	AC068764 Homo sapi
125	25	0.9	100000	9	AP000183	AP000183 Homo sapi	C 198	24	0.9	185832	9	AL137800	AL137800 Human DNA
126	25	0.9	105062	2	AC074049	AC074049 Homo sapi	C 199	24	0.9	186138	2	AC093286	AC093286 Homo sapi
127	25	0.9	123818	9	HS105613	AL031727 Human DNA	C 200	24	0.9	186629	2	AL512603	AL512603 Homo sapi
128	25	0.9	123925	9	AC003976	AC003976 Homo sapi	201	24	0.9	189169	2	AC0933008	AC0933008 Homo sapi
129	25	0.9	127587	2	AC005959	AC005959 Homo sapi	C 202	24	0.9	192550	9	AC009779	AC009779 Homo sapi
130	25	0.9	129098	9	HS46H23	284443 Human DNA s	203	24	0.9	192722	2	AL590103	AL590103 Homo sapi
131	25	0.9	130080	9	AP003160	AC107976 Homo sapi	C 204	24	0.9	200562	2	AC024087	AC024087 Homo sapi
132	25	0.9	136690	2	AC107976	AC068859 Homo sapi	C 205	24	0.9	205199	2	AC090573	AC090573 Homo sapi
133	25	0.9	143996	2	AC068859	AL137789 Human DNA	C 206	24	0.9	208659	2	AC018675	AC018675 Homo sapi
134	25	0.9	145325	9	AL137789	AC011455 Homo sapi	C 207	24	0.9	215861	9	HS455J7	HS455J7 Homo sapi
135	25	0.9	148876	9	AC011455	AC011455 Homo sapi	C 208	24	0.9	228045	2	AC093298	AC093298 Homo sapi
136	25	0.9	150829	2	AC009092	292844 Human DNA s	C 209	24	0.9	231247	9	AC008543	AC008543 Homo sapi
137	25	0.9	151798	9	HS435C23	AC004674 Homo sapi	210	24	0.9	340000	9	HS21C049	HS21C049 Homo sapi
138	25	0.9	158866	9	AC004674	AP002962 Homo sapi	C 211	23	0.8	261	11	G34057	G34057 human STS S
139	25	0.9	160222	9	AP002962	AP002962 Homo sapi	212	23	0.8	305	11	G34057	G34057 human STS S
140	25	0.9	161757	2	AC106000	AL360227 Human DNA	213	23	0.8	2103	10	MUSZFP51	D16630 Mus musculu
141	25	0.9	161938	9	AL360227	AL360227 Human DNA	214	23	0.8	2141	10	BC010591	BC010591 Mus muscu
142	25	0.9	162606	2	AC016348	AL360227 Human DNA	215	23	0.8	2200	10	MMU41671	U041671 Mus musculu
143	25	0.9	166329	2	AC016348	AL139187 Human DNA	216	23	0.8	2643	9	HSU09366	HSU09366 Human zinc
144	25	0.9	166961	2	AC067990	AC067990 Homo sapi	217	23	0.8	2718	9	BC001887	BC001887 Homo sapi
145	25	0.9	173902	2	AC073364	AC073364 Homo sapi	218	23	0.8	2973	10	MMZFPTA	X63747 M. musculus
146	25	0.9	179106	2	AC017059	AC017059 Homo sapi	C 219	23	0.8	48023	9	AL662888	AL662888 Human DNA

220	23	0.8	54594	9	AC073272	AC073272 Homo sapi	293	22	0.8	439	6	AX072687	AX072687 Sequence
221	23	0.8	59270	2	AC087617	AC087617 Homo sapi	294	22	0.8	484	6	AX321745	AX321745 Sequence
222	23	0.8	63037	2	AC090194	AC090194 Homo sapi	295	22	0.8	791	9	HSU20648	U20648 Human zinc
223	23	0.8	68589	2	AC087682	AC087682 Homo sapi	296	22	0.8	1553	10	AF332090	AF332090 Mus muscu
224	23	0.8	70204	2	AC010737	AC010737 Homo sapi	297	22	0.8	1557	10	AF332089	AF332089 Mus muscu
225	23	0.8	71822	2	AC027041	AC027041 Homo sapi	298	22	0.8	1739	9	HSNCF101	U00776 Human neutr
226	23	0.8	74690	2	AC022581	AC022581 Homo sapi	299	22	0.8	2295	10	AF226870	AF226870 Mus muscu
227	23	0.8	7690	2	AC022581	AC022581 Homo sapi	300	22	0.8	2382	9	AF114816	AF114816 Homo sapi
228	23	0.8	78603	2	HS1188A21	HS117373 Human DNA	301	22	0.8	2411	10	AF277902	AF277902 Rattus no
229	23	0.8	81417	2	AC023265	AC023265 Homo sapi	302	22	0.8	2432	9	HSU09413	U09413 Human zinc
230	23	0.8	99035	2	HS90012	AL031906 Human DNA	303	22	0.8	2679	6	AX274918	AX274918 Sequence
231	23	0.8	102507	9	HSJ32AN14	AL109938 Human DNA	304	22	0.8	3031	10	AF167320	AF167320 Mus muscu
232	23	0.8	106233	9	HS43C13	AL009175 Homo sapi	305	22	0.8	3146	9	AF154846	AF154846 Homo sapi
233	23	0.8	110000	2	AL359032-1	Continuation (2 of	306	22	0.8	3271	9	BC007256	BC007256 Homo sapi
234	23	0.8	110271	2	AC024576	AC024576 Homo sapi	307	22	0.8	5541	9	AK000086	AK000086 Homo sapi
235	23	0.8	112878	9	AC006999	AC006999 Homo sapi	308	22	0.8	35641	6	AX332256	AX332256 Sequence
236	23	0.8	115296	2	AC084086	AC084086 Homo sapi	309	22	0.8	35641	6	HUMERGC46	L76568 Homo sapien
237	23	0.8	128444	2	AF214634	AF214634 Homo sapi	310	22	0.8	41601	9	AC004017	AC004017 Homo sapi
238	23	0.8	137557	2	AC005909	AC005909 Homo sapi	311	22	0.8	43958	9	AC021011	AC021011 Homo sapi
239	23	0.8	142544	2	AC108020	AC108020 Homo sapi	312	22	0.8	47466	8	AP002061	AP002061 Arabidops
240	23	0.8	144234	2	AP002425	AP002425 Homo sapi	313	22	0.8	63724	9	AL158081	AL158081 Human DNA
241	23	0.8	144279	2	AP001570	AP001570 Homo sapi	314	22	0.8	84114	9	AC003006	AC003006 Human DNA
242	23	0.8	145722	2	AC015503	AC015503 Homo sapi	315	22	0.8	88509	9	HSJ159G19	AL078462 Human DNA
243	23	0.8	150113	9	AC015988	AC015988 Homo sapi	316	22	0.8	97097	9	AL353740	AL353740 Human DNA
244	23	0.8	155163	30	AC027502	AC027502 Homo sapi	317	22	0.8	103601	9	CNS01DUU	AL133369 Human chr
245	23	0.8	155313	2	AC011091	AC011091 Homo sapi	318	22	0.8	110000	2	AC008749-0	AC008749 Homo sapi
246	23	0.8	155432	9	HSJD568F9	AL049646 Human DNA	319	22	0.8	112186	9	HSJD80L13	AL132981 Human DNA
247	23	0.8	155731	2	AP001205	AP001205 Homo sapi	320	22	0.8	119453	9	AC004006	AC004006 Homo sapi
248	23	0.8	155827	9	CNS05FCT	AL355099 Human chr	321	22	0.8	124154	2	AC093217	AC093217 Homo sapi
249	23	0.8	156997	9	AC011465	AC011465 Homo sapi	322	22	0.8	126795	2	AC010266	AC010266 Homo sapi
250	23	0.8	157564	2	AC025996	AC025996 Homo sapi	323	22	0.8	133418	9	AL353709	AL353709 Human DNA
251	23	0.8	158145	2	AC104664	AC104664 Homo sapi	324	22	0.8	138752	9	AL161627	AL161627 Human DNA
252	23	0.8	158357	2	CNS07EFC	AL512414 Human chr	325	22	0.8	140053	2	AC026745	AC026745 Homo sapi
253	23	0.8	160640	2	AC073494	AC073494 Homo sapi	326	22	0.8	141372	2	AC011779	AC011779 Homo sapi
254	23	0.8	163569	2	AC079295	AC079295 Homo sapi	327	22	0.8	142037	2	AC025171	AC025171 Homo sapi
255	23	0.8	164831	9	AL353765	AL353765 Human DNA	328	22	0.8	151272	2	AC009763	AC009763 Homo sapi
256	23	0.8	166009	2	AC011061	AC011061 Homo sapi	329	22	0.8	151791	2	AC108695	AC108695 Homo sapi
257	23	0.8	166207	9	AL135785	AL135785 Human DNA	330	22	0.8	152050	9	AC079408	AC079408 Homo sapi
258	23	0.8	166966	9	AC016970	AC016970 Homo sapi	331	22	0.8	155837	9	AL445495	AL445495 Human DNA
259	23	0.8	168042	2	AC025931	AC025931 Homo sapi	332	22	0.8	156181	2	AC011677	AC011677 Homo sapi
260	23	0.8	168217	9	AL138920	AL138920 Human DNA	333	22	0.8	157812	2	AC055605	AC055605 Homo sapi
261	23	0.8	168822	9	AC007511	AC007511 Homo sapi	334	22	0.8	159250	9	AC073528	AC073528 Homo sapi
262	23	0.8	171179	2	AC034140	AC034140 Homo sapi	335	22	0.8	159898	2	AC104775	AC104775 Homo sapi
263	23	0.8	171545	9	AC008394	AC008394 Homo sapi	336	22	0.8	159998	2	AC063974	AC063974 Homo sapi
264	23	0.8	174080	2	AF217246	AF217246 Homo sapi	337	22	0.8	160463	2	AC011566	AC011566 Homo sapi
265	23	0.8	176213	9	AC009518	AC009518 Human chr	338	22	0.8	163198	2	AL161634	AL161634 Homo sapi
266	23	0.8	177286	9	AC008085	AC008085 Homo sapi	339	22	0.8	164220	30	AC024130	AC024130 Homo sapi
267	23	0.8	177522	9	AC093566	AC093566 Homo sapi	340	22	0.8	164479	9	AC103922	AC103922 Homo sapi
268	23	0.8	179238	2	AC009451	AC009451 Homo sapi	341	22	0.8	16538	9	AC093805	AC093805 Homo sapi
269	23	0.8	179876	2	AC013277	AC013277 Homo sapi	342	22	0.8	165589	9	AL356584	AL356584 Human DNA
270	23	0.8	182506	9	AL359077	AL359077 Human DNA	343	22	0.8	165608	2	AC018473	AC018473 Homo sapi
271	23	0.8	183599	2	AC073501	AC073501 Homo sapi	344	22	0.8	166983	2	AC002541	AC002541 Human BAC
272	23	0.8	184510	2	AC018894	AC018894 Homo sapi	345	22	0.8	167330	2	AC021822	AC021822 Homo sapi
273	23	0.8	185322	2	AC025667	AC025667 Mus muscu	346	22	0.8	167943	9	HS267W20	AL031053 Human DNA
274	23	0.8	186208	30	AC019204	AC019204 Homo sapi	347	22	0.8	168779	9	AC026405	AC026405 Homo sapi
275	23	0.8	187253	2	AC104186	AC104186 Homo sapi	348	22	0.8	169089	9	AC008751	AC008751 Homo sapi
276	23	0.8	188992	2	AC073873	AC073873 Homo sapi	349	22	0.8	169418	2	AC092046	AC092046 Homo sapi
277	23	0.8	189008	2	AP001491	AP001491 Homo sapi	350	22	0.8	169418	2	AC092046	AC092046 Homo sapi
278	23	0.8	189241	2	AC090402	AC090402 Homo sapi	351	22	0.8	169849	2	AC108715	AC108715 Homo sapi
279	23	0.8	189370	2	AC022254	AC022254 Homo sapi	352	22	0.8	171495	2	AC027571	AC027571 Homo sapi
280	23	0.8	191832	2	AC011104	AC011104 Homo sapi	353	22	0.8	172687	2	AC027144	AC027144 Homo sapi
281	23	0.8	193057	9	AC022325	AC022325 Homo sapi	354	22	0.8	172720	2	AC091886	AC091886 Homo sapi
282	23	0.8	194851	9	AC079264	AC079264 Homo sapi	355	22	0.8	173917	2	AC027476	AC027476 Homo sapi
283	23	0.8	197143	2	AC069173	AC069173 Homo sapi	356	22	0.8	174342	2	AC016901	AC016901 Homo sapi
284	23	0.8	198742	2	AL355360	AL355360 Homo sapi	357	22	0.8	175552	2	AC027475	AC027475 Homo sapi
285	23	0.8	199534	2	AC025959	AC025959 Homo sapi	358	22	0.8	180925	2	AL0603926	AL0603926 Homo sapi
286	23	0.8	200820	2	AC107896	AC107896 Homo sapi	359	22	0.8	181438	2	AC023047	AC023047 Homo sapi
287	23	0.8	201344	2	AC108496	AC108496 Homo sapi	360	22	0.8	183072	2	AC026338	AC026338 Homo sapi
288	23	0.8	205195	2	AC025963	AC025963 Homo sapi	361	22	0.8	183215	2	AC023370	AC023370 Homo sapi
289	23	0.8	207757	2	AL589784	AL589784 Homo sapi	362	22	0.8	183228	9	AC010547	AC010547 Homo sapi
290	23	0.8	209958	2	AC087135	AC087135 Mus muscu	363	22	0.8	184181	9	AL158141	AL158141 Human DNA
291	22	0.8	282	9	HMZNFEN	HM8370 Homo sapien	364	22	0.8	185322	2	AC025667	AC025667 Mus muscu
292	22	0.8	292	9	HSZNFPT3	MX5231 H. sapiens m	365	22	0.8	185622	9	AC091667	AC091667 Homo sapi

366	22	0.8	187628	9	AC011311	Homo sapi	C 439	21	0.8	140591	2	AC090888	AC090888	Homo sapi
367	22	0.8	188629	9	AC084365	Homo sapi	C 440	21	0.8	141441	9	AC087073	AC087073	Homo sapi
368	22	0.8	190499	2	AC068392	Homo sapi	C 441	21	0.8	143794	9	AC009315	AC009315	Homo sapi
369	22	0.8	190907	9	AC005261	Homo sapi	C 442	21	0.8	144374	2	AC009283	AC009283	Homo sapi
370	22	0.8	191960	9	CNS01R68	Human chr	C 443	21	0.8	145299	2	AF328728	AF328728	Homo sapi
371	22	0.8	192529	2	AC019190	Homo sapi	C 444	21	0.8	145408	2	AC024303	AC024303	Homo sapi
372	22	0.8	193041	9	AC092017	Homo sapi	C 445	21	0.8	147031	2	AC084264	AC084264	Homo sapi
373	22	0.8	195358	2	AC040892	Homo sapi	C 446	21	0.8	151202	9	AL160168	AL160168	Human DNA
374	22	0.8	195653	9	AC005873	ctb_15_o	C 447	21	0.8	153704	9	AC008567	AC008567	Homo sapi
375	22	0.8	196319	9	CNS01DR4	Human chr	C 448	21	0.8	154600	2	AC110610	AC110610	Homo sapi
376	22	0.8	198751	9	AC022148	Homo sapi	C 449	21	0.8	154917	9	AC026185	AC026185	Homo sapi
377	22	0.8	202077	2	AC073911	Homo sapi	C 450	21	0.8	155459	2	AC016619	AC016619	Homo sapi
378	22	0.8	204924	2	AC015616	Homo sapi	C 451	21	0.8	155359	2	AC026477	AC026477	Homo sapi
379	22	0.8	205815	2	AC008759	Homo sapi	C 452	21	0.8	156045	2	AC026710	AC026710	Homo sapi
380	22	0.8	206346	2	AC027438	Homo sapi	C 453	21	0.8	156279	2	AC027707	AC027707	Homo sapi
381	22	0.8	209958	2	AC087135	Mus muscu	C 454	21	0.8	156681	2	AC104818	AC104818	Homo sapi
382	22	0.8	211259	2	AC024928	Homo sapi	C 455	21	0.8	157300	10	AL592522	AL592522	Mouse DNA
383	22	0.8	218233	2	AC099669	Homo sapi	C 456	21	0.8	157349	2	AF145206	AF145206	Homo sapi
384	22	0.8	230864	2	AC079523	Mus muscu	C 457	21	0.8	158213	2	AF214635	AF214635	Homo sapi
385	22	0.8	231146	2	AL669941	Mus muscu	C 458	21	0.8	158724	9	AL160010	AL160010	Human DNA
386	22	0.8	236029	2	AC019232	Homo sapi	C 459	21	0.8	160182	2	AC026594	AC026594	Homo sapi
387	22	0.8	245235	2	AL611964	Homo sapi	C 460	21	0.8	160334	2	AC079549	AC079549	Homo sapi
388	21	0.8	392	9	HUMNR70H10	Homo sapi	C 461	21	0.8	160615	9	AC026786	AC026786	Homo sapi
389	21	0.8	444	9	HSZNF19S1	Human chr	C 462	21	0.8	161078	9	AP003350	AP003350	Homo sapi
390	21	0.8	538	5	EMZNF6	Human chr	C 463	21	0.8	161286	2	AC025120	AC025120	Homo sapi
391	21	0.8	538	5	EMZNF6M	Human chr	C 464	21	0.8	161649	2	AC023652	AC023652	Homo sapi
392	21	0.8	1333	9	BC014381	Homo sapi	C 465	21	0.8	162700	2	AC024105	AC024105	Homo sapi
393	21	0.8	1414	9	HSK0X1	Human chr	C 466	21	0.8	162943	9	AL355612	AL355612	Human chr
394	21	0.8	1493	9	AF193057	Homo sapi	C 467	21	0.8	164374	9	AL355612	AL355612	Human chr
395	21	0.8	1494	9	AF217988	Homo sapi	C 468	21	0.8	164719	2	AC068623	AC068623	Homo sapi
396	21	0.8	1529	9	BC007393	Homo sapi	C 469	21	0.8	165079	2	AC024517	AC024517	Homo sapi
397	21	0.8	1625	9	AK000223	Homo sapi	C 470	21	0.8	165693	2	AC087879	AC087879	Homo sapi
398	21	0.8	1663	9	AF155100	Homo sapi	C 471	21	0.8	165905	2	AC026752	AC026752	Homo sapi
399	21	0.8	1675	9	HSMB00838	Homo sapi	C 472	21	0.8	166862	2	AC069533	AC069533	Homo sapi
400	21	0.8	1960	9	HSMB01833	Homo sapi	C 473	21	0.8	166937	2	AC012255	AC012255	Homo sapi
401	21	0.8	2083	9	AF166122	Homo sapi	C 474	21	0.8	167275	2	AC074123	AC074123	Homo sapi
402	21	0.8	2203	9	AK056313	Homo sapi	C 475	21	0.8	167300	2	AC021394	AC021394	Homo sapi
403	21	0.8	2330	9	BC001791	Homo sapi	C 476	21	0.8	167594	2	AC095035	AC095035	Homo sapi
404	21	0.8	2382	6	AR170700	Sequence	C 477	21	0.8	167806	2	AC092572	AC092572	Homo sapi
405	21	0.8	2566	9	AK056489	Homo sapi	C 478	21	0.8	168440	2	AC105282	AC105282	Homo sapi
406	21	0.8	2732	9	AK027700	Homo sapi	C 479	21	0.8	168474	3	AC093454	AC093454	Drosophila
407	21	0.8	2864	9	AK056548	Homo sapi	C 480	21	0.8	169117	9	AC087894	AC087894	Homo sapi
408	21	0.8	3007	10	BC021326	Mus muscu	C 481	21	0.8	169333	2	AC108119	AC108119	Homo sapi
409	21	0.8	3059	9	BC006436	Homo sapi	C 482	21	0.8	169740	2	AF186193	AF186193	Homo sapi
410	21	0.8	3204	9	AK027363	Homo sapi	C 483	21	0.8	169944	4	AC091643	AC091643	Homo sapi
411	21	0.8	11000	8	AF286114	Eremothec	C 484	21	0.8	170445	2	AC068594	AC068594	Homo sapi
412	21	0.8	42967	9	HUMPCBD	Human DNA	C 485	21	0.8	170656	2	AP001263	AP001263	Homo sapi
413	21	0.8	42967	9	AC093619	Homo sapi	C 486	21	0.8	171012	9	AC073236	AC073236	Homo sapi
414	21	0.8	44059	9	AL669815	Human DNA	C 487	21	0.8	171548	2	AC024732	AC024732	Homo sapi
415	21	0.8	52346	2	AC084346	Homo sapi	C 488	21	0.8	172421	2	AC069538	AC069538	Homo sapi
416	21	0.8	56745	2	AC084347	Homo sapi	C 489	21	0.8	173118	2	AC078827	AC078827	Homo sapi
417	21	0.8	59228	2	AC069452	Homo sapi	C 490	21	0.8	173746	2	AC024040	AC024040	Homo sapi
418	21	0.8	74709	2	AC105479	Rattus no	C 491	21	0.8	174612	9	AC018494	AC018494	Homo sapi
419	21	0.8	79461	2	AC016422	Homo sapi	C 492	21	0.8	174831	2	AC089666	AC089666	Homo sapi
420	21	0.8	84873	2	AC024570	Homo sapi	C 493	21	0.8	175256	9	AC090775	AC090775	Homo sapi
421	21	0.8	95738	2	AC020272	Drosophila	C 494	21	0.8	175339	2	AC005772	AC005772	Homo sapi
422	21	0.8	97959	2	AL590863	Homo sapi	C 495	21	0.8	175454	2	AC073176	AC073176	Homo sapi
423	21	0.8	98889	2	AL590863	Homo sapi	C 496	21	0.8	175529	2	AL591916	AL591916	Homo sapi
424	21	0.8	103054	2	AC014803	Drosophila	C 497	21	0.8	176018	9	AC093878	AC093878	Homo sapi
425	21	0.8	108892	2	AC034232	Homo sapi	C 498	21	0.8	176197	2	AC092043	AC092043	Homo sapi
426	21	0.8	111837	9	AC069383	Homo sapi	C 499	21	0.8	177267	9	AC026116	AC026116	Homo sapi
427	21	0.8	112359	9	HSU60822	Human chr	C 500	21	0.8	177636	2	AC094891	AC094891	Homo sapi
428	21	0.8	113345	9	AF109076	Homo sapi	C 501	21	0.8	177834	9	AC018811	AC018811	Homo sapi
429	21	0.8	113654	9	AL606526	Human DNA	C 502	21	0.8	178141	2	AC074345	AC074345	Homo sapi
430	21	0.8	123530	9	AL139110	Human DNA	C 503	21	0.8	178184	9	AC008770	AC008770	Homo sapi
431	21	0.8	128386	2	AL356782	Human DNA	C 504	21	0.8	178246	2	AP003782	AP003782	Homo sapi
432	21	0.8	132390	2	AC008899	Homo sapi	C 505	21	0.8	178367	9	AC006115	AC006115	Homo sapi
433	21	0.8	134249	9	AC019116	Homo sapi	C 506	21	0.8	178977	9	AC092505	AC092505	Homo sapi
434	21	0.8	135345	9	AL356294	Human DNA	C 507	21	0.8	179319	2	AC109493	AC109493	Homo sapi
435	21	0.8	138076	2	AC099169	Rattus no	C 508	21	0.8	179581	2	AC109397	AC109397	Homo sapi
436	21	0.8	139476	9	AC008773	Homo sapi	C 509	21	0.8	180510	2	AC022415	AC022415	Homo sapi
437	21	0.8	139623	2	AC040932	Homo sapi	C 510	21	0.8	180615	9	AC025549	AC025549	Homo sapi
438	21	0.8	140156	2	AC027261	Homo sapi	C 511	21	0.8	180672	9	AC011451	AC011451	Homo sapi

512	21	0.8	180717	2	AC068169	AC068169	Homo sapi	585	20	0.7	2627	9	BC013741	BC013741	Homo sapi
513	21	0.8	181141	2	AC093621	AC093621	Homo sapi	586	20	0.7	2640	10	MMNRVZFP	X74855 M.musculus	
514	21	0.8	181312	2	AC091117	AC091117	Homo sapi	587	20	0.7	2779	9	HSMB00634	AL096747 Homo sapi	
515	21	0.8	181720	3	AC010114	AC010114	Drosophila	588	20	0.7	2821	10	RNU56732	U56732 Rattus norv	
516	21	0.8	181858	2	AC084856	AC084856	Homo sapi	589	20	0.7	2871	9	HSR245587	AJ745587 Homo sapi	
517	21	0.8	184043	2	AC092835	AC092835	Homo sapi	590	20	0.7	2973	5	XISTAR	X84996 X.laeyvis mR	
518	21	0.8	185091	2	AC013835	AC013835	Homo sapi	591	20	0.7	3759	9	HS2NF742	X52715 H.sapiens m	
519	21	0.8	185916	2	AC021646	AC021646	Homo sapi	592	20	0.7	4582	10	MUSPLM24X	M85502 Mus musculus	
520	21	0.8	186024	2	AC084225	AC084225	Homo sapi	593	20	0.7	9367	9	HSBOK2EX2	X81126 H.sapiens H	
521	21	0.8	186091	2	AC011242	AC011242	Homo sapi	594	20	0.7	25103	2	AC107199	AC107199 Rattus no	
522	21	0.8	186265	2	AC078959	AC078959	Homo sapi	595	20	0.7	31938	2	AC094656	AC094656 Rattus no	
523	21	0.8	186793	3	AC010660	AC010660	Drosophila	596	20	0.7	35715	9	AC020951	AC020951 Homo sapi	
524	21	0.8	186868	2	AC023639	AC023639	Homo sapi	597	20	0.7	36510	9	AC004397	AC004397 Homo sapi	
525	21	0.8	187415	2	AC026352	AC026352	Homo sapi	598	20	0.7	39526	9	AC003961	AC003961 Human Cos	
526	21	0.8	187746	2	AC087123	AC087123	Mus muscu	599	20	0.7	42861	3	CEFS2B11	Z82268 Caenorhabdi	
527	21	0.8	188734	2	AC069443	AC069443	Homo sapi	600	20	0.7	50396	2	AC105042	AC105042 Homo sapi	
528	21	0.8	190268	2	AC108222	AC108222	Homo sapi	601	20	0.7	51099	2	AC096466_3	Continuation (4 of	
529	21	0.8	192963	9	AL354763	AL354763	Human DNA	602	20	0.7	55823	2	AC040923	AC040923 Homo sapi	
530	21	0.8	193935	2	AC020609	AC020609	Homo sapi	603	20	0.7	57145	9	AL391820	AL391820 Human DNA	
531	21	0.8	195280	2	AC025081	AC025081	Homo sapi	604	20	0.7	57287	9	AC008283	AC008283 Homo sapi	
532	21	0.8	197425	9	AC016776	AC016776	Homo sapi	605	20	0.7	57594	2	AC104552	AC104552 Mus muscu	
533	21	0.8	198586	2	AC092503	AC092503	Homo sapi	606	20	0.7	59712	2	AC104221	AC104221 Mus muscu	
534	21	0.8	201437	2	AC015981	AC015981	Homo sapi	607	20	0.7	62716	2	AC012201	AC012201 Homo sapi	
535	21	0.8	203006	9	HS1145L23	HS1145L23	Human DNA	608	20	0.7	64008	2	AC105101	AC105101 Homo sapi	
536	21	0.8	207908	2	AC108668	AC108668	Homo sapi	609	20	0.7	64365	2	AC099842	AC099842 Homo sapi	
537	21	0.8	218233	2	AC099669	AC099669	Homo sapi	610	20	0.7	66975	9	AL135791	AL135791 Human DNA	
538	21	0.8	218445	2	AC083959	AC083959	Homo sapi	611	20	0.7	73198	9	HSB127L3	HSB127L3 Homo sapi	
539	21	0.8	218670	9	AC096669	AC096669	Homo sapi	612	20	0.7	74105	2	AC087648	AC087648 Homo sapi	
540	21	0.8	218977	9	AC007500	AC007500	Homo sapi	613	20	0.7	74645	2	AC009592	AC009592 Homo sapi	
541	21	0.8	219658	9	AC022681	AC022681	Homo sapi	614	20	0.7	78667	2	AC023522	AC023522 Homo sapi	
542	21	0.8	223791	2	AC084042	AC084042	Mus muscu	615	20	0.7	86719	9	HSJ906P16	HSJ906P16 Human DNA	
543	21	0.8	223805	2	AC079475	AC079475	Mus muscu	616	20	0.7	87903	9	HSJ37B17	HSJ37B17 Human DNA	
544	21	0.8	226116	2	AC084078	AC084078	Homo sapi	617	20	0.7	93643	9	HS629A15	AL033530 Human DNA	
545	21	0.8	226860	2	AL359920	AL359920	Human DNA	618	20	0.7	103367	9	HSB23P16	AL036641 Human DNA	
546	21	0.8	229687	2	AC107068	AC107068	Homo sapi	619	20	0.7	105412	2	AC011458	AC011458 Homo sapi	
547	21	0.8	241714	2	AC083834	AC083834	Mus muscu	620	20	0.7	106341	9	AL390316	AL390316 Human DNA	
548	21	0.8	248940	2	AC103563	AC103563	Homo sapi	621	20	0.7	108542	9	AC104562	AC104562 Homo sapi	
549	21	0.8	250681	2	AP004282	AP004282	Homo sapi	622	20	0.7	108542	2	AF263550	AF263550 Homo sapi	
550	21	0.8	257383	2	AC079494	AC079494	Mus muscu	623	20	0.7	110000	2	AL135900	AL135900 Homo sapi	
551	21	0.8	274656	2	AC079538	AC079538	Mus muscu	624	20	0.7	110000	2	AL596328_3	Continuation (4 of	
552	21	0.8	277603	2	AC079543	AC079543	Mus muscu	625	20	0.7	113022	2	AC084846	AC084846 Homo sapi	
553	21	0.8	283821	3	ABE03528	ABE03528	Drosophila	626	20	0.7	113284	2	AC074300	AC074300 Homo sapi	
554	21	0.8	344150	9	AF235103	AF235103	Homo sapi	627	20	0.7	119034	9	AC007544	AC007544 Homo sapi	
555	20	0.7	99	10	RNU78142	RNU78142	Rattus norv	628	20	0.7	120450	9	AC093393	AC093393 Homo sapi	
556	20	0.7	413	6	AX072834	AX072834	Sequence	629	20	0.7	120799	2	AC092120	AC092120 Homo sapi	
557	20	0.7	426	11	G58442	G58442	SHGC-104386	630	20	0.7	127058	2	AC025652	AC025652 Homo sapi	
558	20	0.7	429	11	G33231	G33231	human STS S	631	20	0.7	127606	2	AC110073	AC110073 Homo sapi	
559	20	0.7	888	6	AX188154	AX188154	Sequence	632	20	0.7	128779	9	HS29X1	Z98745 Human DNA	
560	20	0.7	1094	9	HS2NF74A2	HS2NF74A2	Homo sapi	633	20	0.7	130067	9	AC007228	AC007228 Homo sapi	
561	20	0.7	1106	5	AF026084	AF026084	Taeniolopy	634	20	0.7	131215	9	AC079614	AC079614 Homo sapi	
562	20	0.7	1263	9	BC022246	BC022246	Homo sapi	635	20	0.7	132442	9	AC069205	AC069205 Homo sapi	
563	20	0.7	1674	9	AK054647	AK054647	Homo sapi	636	20	0.7	135432	2	AC069036	AC069036 Homo sapi	
564	20	0.7	1699	9	BC007868	BC007868	Homo sapi	637	20	0.7	136492	2	AC011143	AC011143 Homo sapi	
565	20	0.7	1776	9	AF242519	AF242519	Homo sapi	638	20	0.7	136725	9	AC026369	AC026369 Homo sapi	
566	20	0.7	1800	5	XLX5NA	XLX5NA	Xenopus lae	639	20	0.7	141579	2	AC044793	AC044793 Homo sapi	
567	20	0.7	1892	6	AR116036	AR116036	Sequence	640	20	0.7	142000	9	AC095353	AC095353 Homo sapi	
568	20	0.7	2080	10	BC010442	BC010442	Mus muscu	641	20	0.7	142796	2	AC006411	AC006411 Homo sapi	
569	20	0.7	2088	3	AB057740	AB057740	Ciona sav	642	20	0.7	146471	2	AC026545	AC026545 Homo sapi	
570	20	0.7	2181	6	AX274847	AX274847	Sequence	643	20	0.7	146768	9	AC016194	AC016194 Homo sapi	
571	20	0.7	2250	10	BC006587	BC006587	Mus muscu	644	20	0.7	147185	9	AC020589	AC020589 Homo sapi	
572	20	0.7	2356	10	MOSK1ZF	MOSK1ZF	Mus muscu	645	20	0.7	147706	9	AC092578	AC092578 Homo sapi	
573	20	0.7	2382	10	MMY14295	MMY14295	Mus muscu	646	20	0.7	147876	9	AC024600	AC024600 Homo sapi	
574	20	0.7	2386	9	HSBOK2H1	HSBOK2H1	H.sapiens H	647	20	0.7	147929	9	AC016951	AC016951 Homo sapi	
575	20	0.7	2407	9	AB066544	AB066544	Macaca fa	648	20	0.7	147978	9	AL136370	AL136370 Human DNA	
576	20	0.7	2416	9	HS2NF741	HS2NF741	H.sapiens z	649	20	0.7	148359	9	AC007434	AC007434 Homo sapi	
577	20	0.7	2416	11	G27154	G27154	SHGC-31560	650	20	0.7	148466	2	AC023167	AC023167 Mus muscu	
578	20	0.7	2454	10	AF184111	AF184111	Mus muscu	651	20	0.7	151777	2	AL354872	AL354872 Human DNA	
579	20	0.7	2478	9	AK027592	AK027592	Homo sapi	652	20	0.7	152144	2	AC016293	AC016293 Homo sapi	
580	20	0.7	2504	9	HS009367	HS009367	Human zinc	653	20	0.7	152144	2	AC016293	AC016293 Homo sapi	
581	20	0.7	2530	9	BC006421	BC006421	Homo sapi	654	20	0.7	152334	2	AL356299	AL356299 Human DNA	
582	20	0.7	2554	10	BC011183	BC011183	Mus muscu	655	20	0.7	152649	2	AC016353	AC016353 Homo sapi	
583	20	0.7	2593	10	AB010143	AB010143	Mus muscu	656	20	0.7	153241	2	AC024960	AC024960 Homo sapi	
584	20	0.7	2604	10	MDZFP30	MDZFP30	Rattus no	657	20	0.7	153479	2	AC095130	AC095130 Rattus no	

658	20	0.7	154307	2	AC108150	Homo sapi	731	20	0.7	186142	2	AL390124	AL390124	Homo sapi
c 659	20	0.7	155094	2	AL512637	Homo sapi	732	20	0.7	186760	2	AC092123	AC092123	Homo sapi
660	20	0.7	156244	9	AC011815	Homo sapi	c 733	20	0.7	187295	2	AC021531	AC021531	Homo sapi
c 661	20	0.7	156466	9	AL354796	Human DNA	c 734	20	0.7	187543	9	AC005678	AC005678	Homo sapi
662	20	0.7	156593	2	AC078999	Homo sapi	c 735	20	0.7	187568	9	AP002967	AP002967	Homo sapi
c 663	20	0.7	158146	2	AC108862	Homo sapi	736	20	0.7	188030	2	AC015456	AC015456	Homo sapi
c 664	20	0.7	158152	2	AC095177	Rattus no	737	20	0.7	188377	2	AC024894	AC024894	Homo sapi
c 665	20	0.7	158370	2	AC097167	Rattus no	738	20	0.7	190000	2	AC027701	AC027701	Homo sapi
c 666	20	0.7	158460	2	AC098177	Rattus no	739	20	0.7	190200	2	AC022021	AC022021	Homo sapi
c 667	20	0.7	158990	2	AL591702	AL591702 Rattus no	c 740	20	0.7	190626	2	AC062036	AC062036	Homo sapi
668	20	0.7	159451	9	AC012618	Homo sapi	c 741	20	0.7	190946	9	CNS01DV1	CNS01DV1	Homo sapi
c 669	20	0.7	159784	2	AC061996	Homo sapi	742	20	0.7	192592	2	AC005500	AC005500	Homo sapi
c 670	20	0.7	159960	2	AC098620	Rattus no	c 743	20	0.7	194052	2	AC053537	AC053537	Homo sapi
671	20	0.7	160003	2	AC021051	AC021051 Homo sapi	744	20	0.7	194297	9	AC044810	AC044810	Homo sapi
672	20	0.7	160242	2	AP002778	AP002778 Homo sapi	c 745	20	0.7	194599	2	AC073668	AC073668	Homo sapi
673	20	0.7	160836	2	AC096630	AC096630 Pan trogl	c 746	20	0.7	195156	9	AC090660	AC090660	Homo sapi
674	20	0.7	162625	2	CEN0883	Z83124 Caenorhabd	c 747	20	0.7	195446	9	AC023170	AC023170	Homo sapi
c 675	20	0.7	165486	2	AC010830	AC010830 Homo sapi	c 748	20	0.7	197328	9	AC010967	AC010967	Homo sapi
676	20	0.7	165566	2	AC092295	AC092295 Homo sapi	749	20	0.7	197784	9	AL450326	AL450326	Homo sapi
677	20	0.7	166288	2	AL627095	AL627095 Homo sapi	c 750	20	0.7	197784	9	AL450326	AL450326	Human DNA
678	20	0.7	167344	2	AC096481	AC096481 Rattus no	751	20	0.7	197819	2	AC073675	AC073675	Homo sapi
c 679	20	0.7	167904	2	AL359173	AL359173 Human DNA	752	20	0.7	198127	2	AP001593	AP001593	Homo sapi
c 680	20	0.7	168870	2	AC021873	AC021873 Homo sapi	753	20	0.7	198440	2	AC009993	AC009993	Homo sapi
c 681	20	0.7	168907	2	AC018390	AC018390 Homo sapi	c 754	20	0.7	198570	9	HS343C1	HS343C1	Homo sapi
c 682	20	0.7	170281	2	AC079910	AC079910 Homo sapi	755	20	0.7	198669	9	AC092378	AC092378	Homo sapi
683	20	0.7	170520	2	AL358172	AL358172 Homo sapi	756	20	0.7	200418	9	AC074138	AC074138	Homo sapi
684	20	0.7	170579	2	AC020985	AC020985 Homo sapi	757	20	0.7	200967	2	AC098511	AC098511	Homo sapi
c 685	20	0.7	170716	2	AC095000	AC095000 Rattus no	758	20	0.7	200991	2	AC106877	AC106877	Homo sapi
c 686	20	0.7	171428	2	AC013407	AC013407 Homo sapi	c 759	20	0.7	201371	2	AC073786	AC073786	Homo sapi
c 687	20	0.7	171779	9	AC079308	AC079308 Homo sapi	c 760	20	0.7	201495	2	AC094415	AC094415	Homo sapi
c 688	20	0.7	172350	2	AC079234	AC079234 Homo sapi	c 761	20	0.7	201895	2	AC097710	AC097710	Homo sapi
c 689	20	0.7	172533	2	AC002381	AL356320 Homo sapi	c 762	20	0.7	204175	2	AL671854	AL671854	Homo sapi
690	20	0.7	172723	2	AL356320	AL356320 Homo sapi	c 763	20	0.7	205380	2	AC037488	AC037488	Homo sapi
c 591	20	0.7	172923	2	AL160255	AL160255 Human DNA	764	20	0.7	206832	2	AL390728	AL390728	Homo sapi
c 592	20	0.7	173188	2	AC023509	AC023509 Homo sapi	765	20	0.7	208660	2	AC092353	AC092353	Homo sapi
c 593	20	0.7	173701	9	AC018946	AC018946 Homo sapi	c 766	20	0.7	209456	2	AC098875	AC098875	Homo sapi
c 594	20	0.7	174429	2	AC104637	AC104637 Homo sapi	c 767	20	0.7	211463	2	AC016508	AC016508	Homo sapi
c 595	20	0.7	174533	2	AC073718	Mus muscu	c 768	20	0.7	212134	2	AC024037	AC024037	Homo sapi
c 696	20	0.7	175654	2	AC025456	AC025456 Homo sapi	c 769	20	0.7	213660	2	AC073699	AC073699	Homo sapi
c 697	20	0.7	175950	2	AC098952	Rattus no	770	20	0.7	214180	2	AC009098	AC009098	Homo sapi
c 698	20	0.7	176512	9	AC012332	AC012332 Homo sapi	c 771	20	0.7	215584	2	AL390779	AL390779	Homo sapi
699	20	0.7	176759	2	AC037464	AC037464 Homo sapi	c 772	20	0.7	216706	9	AC090420	AC090420	Homo sapi
c 700	20	0.7	177164	2	AC103450	AC103450 Rattus no	c 773	20	0.7	218859	2	AC022912	AC022912	Homo sapi
701	20	0.7	177444	2	AC011468	AC011468 Homo sapi	c 774	20	0.7	219043	2	AC073812	AC073812	Mus muscu
702	20	0.7	178184	9	AC008770	AC008770 Homo sapi	c 775	20	0.7	220678	2	AC023857	AC023857	Homo sapi
703	20	0.7	178323	9	AL354810	AL354810 Human DNA	776	20	0.7	223192	2	AL671899	AL671899	Mus muscu
704	20	0.7	178624	9	AC069020	AC069020 Homo sapi	c 777	20	0.7	224573	2	AC008758	AC008758	Homo sapi
705	20	0.7	178624	9	AC009756	AC009756 Homo sapi	c 778	20	0.7	226178	2	AL669827	AL669827	Mus muscu
706	20	0.7	179272	2	AC069149	AC069149 Homo sapi	c 779	20	0.7	227074	2	AC060772	AC060772	Mus muscu
c 707	20	0.7	179272	2	AC069149	AC069149 Homo sapi	780	20	0.7	229000	2	AL391381	AL391381	Homo sapi
c 708	20	0.7	179380	9	AL355501	AL355501 Human DNA	c 781	20	0.7	241099	2	AC099389	AC099389	Rattus no
709	20	0.7	179491	2	AC020673	AC020673 Homo sapi	c 782	20	0.7	242406	2	AL669845	AL669845	Mus muscu
c 710	20	0.7	179526	2	AC034133	AC034133 Homo sapi	c 783	20	0.7	244931	2	AC079495	AC079495	Mus muscu
711	20	0.7	179581	2	AC009397	AC009397 Homo sapi	c 784	20	0.7	267635	2	AC073772	AC073772	Mus muscu
712	20	0.7	179767	9	AC069223	AC069223 Homo sapi	785	20	0.7	272168	9	AC079561	AC079561	Mus muscu
713	20	0.7	180331	9	AC007375	AC007375 Homo sapi	786	20	0.7	275159	9	U82670	U82670	Homo sapien
714	20	0.7	180510	9	AC022415	AC022415 Homo sapi	787	20	0.7	297900	2	AC006704	AC006704	Caenorhab
c 715	20	0.7	180638	9	AC068763	AC068763 Homo sapi	788	20	0.7	300050	9	HSWX1A	HSWX1A	Homo sapi
c 716	20	0.7	180727	2	AC027221	AC027221 Homo sapi	789	20	0.7	318503	2	AC074166	AC074166	Mus muscu
717	20	0.7	180816	2	AL627249	AL627249 Dantio rer	c 790	20	0.7	318503	2	AC074166	AC074166	Mus muscu
718	20	0.7	180900	2	AC018975	AC018975 Homo sapi	c 791	20	0.7	340000	9	HS21C084	HS21C084	Homo sapi
719	20	0.7	180964	2	AC008955	AC008955 Homo sapi	c 792	20	0.7	308	9	HS261Z9	HS261Z9	H. sapiens
720	20	0.7	181026	2	AL591642	AL591642 Homo sapi	c 793	19	0.7	326	11	G33841	G33841	human STS
721	20	0.7	181421	2	AC011637	AC011637 Homo sapi	794	19	0.7	357	3	AP077675	AP077675	Strongylo
c 722	20	0.7	181451	9	AP003038	AP003038 Homo sapi	795	19	0.7	375	11	G34024	G34024	human STS
723	20	0.7	181709	2	AL672031	Mus muscu	796	19	0.7	375	11	HUM4S1258	HUM4S1258	human chrom
724	20	0.7	181773	2	AL365365	Homo sapi	c 797	19	0.7	378	9	HS16C7004	HS16C7004	H. sapiens t
725	20	0.7	182125	2	AC090105	AC090105 Homo sapi	798	19	0.7	387	11	G26051	G26051	human STS
726	20	0.7	182188	9	AC048353	AC048353 Homo sapi	c 799	19	0.7	397	11	G60159	G60159	SHGC-131044
727	20	0.7	182617	9	AC007731	AC007731 Homo sapi	c 800	19	0.7	412	11	HSPE33D08	HSPE33D08	H. sapiens
c 728	20	0.7	184439	2	AP003085	AP003085 Homo sapi	c 801	19	0.7	462	11	G51998	G51998	SHGC-81978
c 729	20	0.7	185493	2	AC068986	AC068986 Homo sapi	802	19	0.7	483	11	G28077	G28077	human STS
c 730	20	0.7	185682	2	AC008732	AC008732 Homo sapi	803	19	0.7	578	11	G56129	G56129	SHGC-101597

C 950	19	0.7	72044	2	AC026582	Homo sapi
C 951	19	0.7	72045	9	AC092214	Homo sapi
C 952	19	0.7	78167	9	AC092028	Homo sapi
C 953	19	0.7	78396	9	AL590486	Human DNA
C 954	19	0.7	81017	2	AC025010	Homo sapi
C 955	19	0.7	81990	2	AC024372	Homo sapi
C 956	19	0.7	82251	17	AC004506	Homo sapi
C 957	19	0.7	82453	9	AC004558	Homo sapi
C 958	19	0.7	83051	9	AL611962	Human DNA
C 959	19	0.7	84096	2	AC074170	Mus muscu
C 960	19	0.7	84113	9	AC073063	Homo sapi
C 961	19	0.7	84613	9	AL136140	Human sapi
C 962	19	0.7	86348	8	AC066300	Arcabidops
C 963	19	0.7	87552	9	HS217016	
C 964	19	0.7	88212	9	AC016679	Human DNA
C 965	19	0.7	89948	9	AL359734	Homo sapi
C 966	19	0.7	89985	30	HSB57233	Human DNA
C 967	19	0.7	92086	2	AF216668	Homo sapi
C 968	19	0.7	92361	2	AC105766	Homo sapi
C 969	19	0.7	93066	9	HSJ224E15	Human DNA
C 970	19	0.7	93458	9	AC079344	Homo sapi
C 971	19	0.7	93902	9	AC011090	Homo sapi
C 972	19	0.7	94162	9	AL450322	Human DNA
C 973	19	0.7	94359	9	AP000974	Homo sapi
C 974	19	0.7	95274	9	HS65019	Human DNA
C 975	19	0.7	95896	9	AL356384	Human DNA
C 976	19	0.7	95963	2	AC106422	Rattus no
C 977	19	0.7	98415	9	AC073421	Homo sapi
C 978	19	0.7	98948	2	AC090275	Homo sapi
C 979	19	0.7	99370	9	AC005057	Homo sapi
C 980	19	0.7	100000	9	AB020862	Homo sapi
C 981	19	0.7	100296	9	AL590308	Human DNA
C 982	19	0.7	100468	9	AC095065	Homo sapi
C 983	19	0.7	100515	9	AL133289	Human DNA
C 984	19	0.7	100733	9	HS201D7	Human DNA
C 985	19	0.7	101854	9	AC091818	Homo sapi
C 986	19	0.7	102135	2	AF215843	Homo sapi
C 987	19	0.7	102198	9	AL157787	Human DNA
C 988	19	0.7	103933	9	HS684024	Human DNA
C 989	19	0.7	104397	2	AF216670	Homo sapi
C 990	19	0.7	104597	2	HS393P12	AL022578
C 991	19	0.7	104901	9	AC090423	Homo sapi
C 992	19	0.7	106783	2	AC010455	Homo sapi
C 993	19	0.7	107781	2	AC095175	Rattus no
C 994	19	0.7	108162	2	AC020960	Mus muscu
C 995	19	0.7	108245	9	AC005166	Homo sapi
C 996	19	0.7	108265	2	AC097020	Rattus no
C 997	19	0.7	108503	2	AL627424	Human DNA
C 998	19	0.7	108634	2	AC020813	Mus muscu
C 999	19	0.7	108907	2	AL135911	Human DNA
C 1000	19	0.7	109056	2	AC074053	Homo sapi

ALIGNMENTS

RESULT	1				
LOCUS	AB013897				
DEFINITION	AB013897	2772 bp	mRNA	linear	PRI 03-NOV-1999
ACCESSION	AB013897				
VERSION	AB013897.1	GI:6177784			
KEYWORDS	HKR1.				
SOURCE	Homo sapiens leukemia cell	cell_line:CMK86	CDNA	to mRNA.	
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (sites)				
AUTHORS	Oguri,T., Katoh,O., Takahashi,T., Isoe,T., Kuramoto,K., Hirata,S., Yamakido,M. and Watanabe,H.				
TITLE	The Kruppel-type Zinc Finger Family Gene, HKR1, Is Induced in Lung Cancer by Exposure to Platinum Drugs				
JOURNAL	Unpublished (1998)				

REFERENCE	2 (bases 1 to 2772)
AUTHORS	Katoh,O.
TITLE	Direct Submission
JOURNAL	Submitted (13-MAY-1998) Osamu Katoh, Res. Inst. Radiation Biology and Medicine, Hiroshima Univ., Department of Environment and Mutation; Kasumi 1-2-3, Minami-Ku., Hiroshima, Hiroshima 734-8553, Japan (E-mail: katoh@mc.ai.med.hiroshima-u.ac.jp, Tel: +81-82-257-5817, Fax: +81-82-256-7104)
FEATURES	
Source	Location/Qualifiers
	1..2772
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19q12-19q13.1"
	/cell_line="CMK86"
	/cell_type="leukemia cell"
gene	1..2096
	/gene="HKR1"
	<1..2096
CDS	/gene="HKR1"
	/function="transcription factor"
	/note="The HKR1 protein is similar to Kruppel-type zinc finger transcription factor and has Kruppel-associated box (KRAB) -A and -B box at the N-terminus (nucleotides 193-390) followed by 13 C2H2-type zinc finger domain at the C-terminus (1020-2084)."
	/codon_start=3
	/product="HKR1"
	/protein_id="DAA86058.1"
	/db_xref="GI:6177785"
	/translation="GALSWLQPKASLIRSFPHLCILLPDFRLDPPRPAFHLISSETLPFRSKSTOETREKMATGILRAKKEAFARVDVAVFTPOEERWLISPAORTLHREVMLETYNLVSLLEIPSSPKPLIAOLREGEAPRVAACPLDCEPSKEPILSPSCPIIFSSGQCEFSLSKSNLKHORSHTEGKPEVCTECGGRFTPKSLTLTHORTSGEKEFPVCAECGGFPNDKSTLISHORTSGEKEPFCREGGRFRORLPHRKHRAHSGAFVCEGGDGFCAKTLTKHQRHAGGAKHVCCECGGFSRSHLIRHORTSGEKEPYICRKGREGFSRKSNIIRHORTHG"
	polyA_site
	2772
	/note="11 a nucleotides"
BASE COUNT	746 a 671 c 720 g 635 t
ORIGIN	
Query Match	100.0%; Score 2772; DB 9; Length 2772;
Best Local Similarity	100.0%; Prid. No. 0;
Matches 2772:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 caggcgcgtlaagcctggtgtgagaccgggaaagcctccctctlaagctcttccacac 60
Db	1 CAGGGCGGTTAAGCTGGGTGGGACCCGGGAGAGGCTCCTCTTAAGGCTTTCCACAC 60
Oy	61 tctgctcctgttacctgaccttgcggttcaggaatcgcgcggcgtgacccgcgttcac 120
Db	61 TCTGCTCCTTGTACTACTGACTTGGGCTTCAGAGATCCCGCGGCTGCACCCGCTTCACATC 120
Oy	121 tgtctctggaaccttggcctcttcccaagaagagacccaggaagccaggaataatggta 180
Db	121 TGTCTTCTTGAGACTTTGGCCCTTCTCCAGCAAGACGACATCAGGAGACCGAGAAATGGCTA 180
Oy	181 cagggctctcagagaccacaaaagagagcgttcgtgcactcaaggagtctggctgttact 240
Db	181 CAGGGCTCTCTAGAGCCAAAAGAGGGCTTCGTGGCATTTCAAGGATGTGGCTGTGTTACT 240
Oy	241 tcaccagagagatgtaggttgtttagccctgctcagaggaacccctcacagaagagtga 300
Db	241 TCACCCAGAGAGAGTGAAGTGTGTTGAGCCCTGCTCAAGAGACCCCTGCACAGGGAGGTGA 300

Qy	301	tgcggagagactttaaaccctcgtgtcctaactgtgaatctcaattcttctaaaccaactca	360
Db	301	tgcggagagactttaaaccctcgtgtcctaactgtgaatctcaattcttctaaaccaactca	360
Qy	361	ttgtcagcctcgagcgagggagagcgccccttgagagagagagagaaatgctcacctggacc	420
Db	361	ttgtctacgcttggagcgagggagagcccctctggagagagagagaaatgctcactggacc	420
Qy	421	tcgttcagaaatcgaagccaagaatccaacttagtcctcctcctgcctctgtatttctca	480
Db	421	tcgttcagaaatcgaagccaagaatccaacttagtcctcctcctgcctctgtatttctca	480
Qy	481	gttcagaaagcttcagccaacatgtgtgtgtcgtatgacatcctcctcaagctgtttccaagt	540
Db	481	gttcagaaagcttcagccaacatgtgtgtgtcgtatgacatcctcctcaagctgtttccaagt	540
Qy	541	tatggcaggaanaatcctctcaccctgvggaanaacatcatalccagaaga tcaagaacaacagc	600
Db	541	tatggcaggaanaatcctctcaccctgvggaanaacatcatalccagaaga tcaagaacaacagc	600
Qy	601	aggatcattctcgtttagtgggaagaacgaatgtatccaagaaggagaaagacccaagac	660
Db	601	aggatcattctcgtttagtgggaagaacgaatgtatccaagaaggagaaagacccaagac	660
Qy	661	tcctgtttggagagatgaagcaaaaatgtgacctcaaaaggaccttcagcccactgaag	720
Db	661	tcctgtttggagagatgaagcaaaaatgtgacctcaaaaggaccttcagcccactgaag	720
Qy	721	aacaaacgccagacacagtcctcaaggaagaaacaacacagctgtgtatataaggtccacccctg	780
Db	721	aacaaacgccagacacagtcctcaaggaagaaacaacacagctgtgtatataaggtccacccctg	780
Qy	781	aacgagagcgagatctcagaagaaacagacaagaatctgcatgttctagaagtctcaagat	840
Db	781	aacgagagcgagatctcagaagaaacagacaagaatctgcatgttctagaagtctcaagat	840
Qy	841	ttggagaatatcaaatatgaaagtttgggccagagctttatccaaggagtccaactcctta	900
Db	841	ttggagaatatcaaatatgaaagtttgggccagagctttatccaaggagtccaactcctta	900
Qy	901	gcctccagaagacaacaacatcggggagagacacactacatgtacatcgtggtggagagacagct	960
Db	901	gcctccagaagacaacaacatcggggagagacacactacatgtacatcgtggtggagagacagct	960
Qy	961	ttggcagtatgtcagtctcctcaatacaaaaacccaagacacacctctgvgggaaagacctatg	1020
Db	961	ttggcagtatgtcagtctcctcaatacaaaaacccaagacacacctctgvgggaaagacctatg	1020
Qy	1021	ttgtcagggagatgtggcgcgagcctttagctgtgaagtccaacctgtgtccacatccagaga	1080
Db	1021	ttgtcagggagatgtggcgcgagcctttagctgtgaagtccaacctgtgtccacatccagaga	1080
Qy	1081	cacacccaagggagaaaacctatgtgtcgaagatgtgtgacagagcttacttggaaagt	1140
Db	1081	cacacccaagggagaaaacctatgtgtcgaagatgtgtgacagagcttacttggaaagt	1140
Qy	1141	cgaaaccttttaacatcagcggagacacttcagggctcaagccttatgtgtgtaaggaat	1200
Db	1141	cgaaaccttttaacatcagcggagacacttcagggctcaagccttatgtgtgtaaggaat	1200
Qy	1201	gtggagcagagctttagacctgaagtccaacctcatcctccacagagggcgacacatgggg	1260
Db	1201	gtggagcagagctttagacctgaagtccaacctcatcctccacagagggcgacacatgggg	1260
Qy	1261	agaaagcctatgtttgcagggaaatgtggcggtgtgtcttcgcagcatccaactgtgtca	1320
Db	1261	agaaagcctatgtttgcagggaaatgtggcggtgtgtcttcgcagcatccaactgtgtca	1320
Qy	1321	gacacaaagagacacatcttcagagagaaagccttaacttcagggagagctgtgagaaagct	1380
Db	1321	gacacaaagagacacatcttcagagagaaagccttaacttcagggagagctgtgagaaagct	1380
Qy	1381	ttagccagaaagctcacactcatcaagcacttlaagagcacacacagagagaaagccttatg	1440

Db	1381	TTAGCCAGAGAGTACACCTTCATCAGACACTTAAAGACACACAGAGAGAGACCTTATG	1440
Qy	1441	tatgcacagaatctgaggcgctcaactttagctgynaatccaactcaaaacacacagagga	1500
Db	1441	TATGCACAGAAATGAGGCGCAGCTTATAGCTGGAAATTCAAACCTCAAAACACACAGAGGA	1500
Qy	1501	caacatcaagggttaaaccttatgtctgcctggaatgagggcagtgctttagccttgaat	1560
Db	1501	CACACTCAGGGGTTTAAACCTTATGTCTGCTGGAGTGGGGCAGTGGCTTTTAGCTCGAAGT	1560
Qy	1561	caaaccttcaacaacacccagaggtcacacacggggagaaagccattgtaatgacgaat	1620
Db	1561	CAAAACCTTACAAACACACAGAGGTCAACACGGGGGAGAACCATTTGTATGTAGCGAGT	1620
Qy	1621	gtgggagaggctttaccggaatcaaccctgatacgcacccagagagacacatcgaagg	1680
Db	1621	GTGGGCGAGGCTTTACCCGGAAATTCAAACCTTATCAGCCACACAGAGGACACATCAGAGG	1680
Qy	1681	agaagcattctgtaatgctgaatgtagacggaggtcttaatgataagtccacccctaat	1740
Db	1681	AGAAGCATTTGTATGTGCTGATGTGGACACGAGCCTTTATATATATATAGTCACCTCATTT	1740
Qy	1741	caacacagagagacacatcaaggagaaagcccttataatgtaaggaagtgtgcsgaagt	1800
Db	1741	CACACCCAGAGAGACACATTCAGGGGAAACCTTTTATGTGCGAGGAGTGTGGCAGAGAGT	1800
Qy	1801	ctggcagaagccttaacctgtttaggcacaagaaggcaaaccaagtgccttcttgta	1860
Db	1801	TTGCGCAGAAAGCCTTAACCTGTTTAGGCACAAAGGGGCACCTCAAGTGCCTTTGTGTGCA	1860
Qy	1861	ggagatgtggagcaaggccttctgtcctaagttaacttcataaacaacagagagacacag	1920
Db	1861	GGGAGTGTGGGCAAGGCTTTTGCTTAAGTAACTCTATTAAACACACAGAGACACACAG	1920
Qy	1921	caaggaggagagcctcaatgtgtgcagaggagtgtggcagaagcctttagccggaggtcaacc	1980
Db	1921	CAGGGGGAAGCCTCATGTGTGCAAGGAGGTGTGGCAAGGCTTTTAGCCGCAAGTCACAC	1980
Qy	1981	tcattagacacccagagagacatctcagagagagaagccttataatctgcagaagtgtagac	2040
Db	1981	TCATTAGACACCCAGAGAGCACATTCAGAGAGAGGCTTATTTTGGAGAAAGTGTGAC	2040
Qy	2041	ggggccttagtcggaagtcacaacctataagacataagaggacaaocccagatagaaac	2100
Db	2041	GGGGCTTTAGTCGGAAATCCAACTTATAGACATCAGAGGACACATCAGATATGAAC	2100
Qy	2101	cttaatgtataagynaaatgtgtacagcctttagccagaagtcatacttcaacagacc	2160
Db	2101	TTTATGTATAGGGAATGTGTACAGACCTTTTATGCCAGAGTCATACTTCATCAGACACCC	2160
Qy	2161	agagagacaacacagatgcctgtggccttttcaagccattgtctagataccaaagtgaagacat	2220
Db	2161	AGAGAGACACACACAGTCTGTGGCTTTTTCAGCCATTCGTTAATACCAAAAGGAGACAT	2220
Qy	2221	ctctgtgtataatgaatgaatgaagccttactgtgaagaactctgatactcaaccctgaag	2280
Db	2221	TCTGTGTCTATTTATGCATATAGACTGTACTGTGAACCTTGTATCTCCATCCACTGAAG	2280
Qy	2281	gagaattgtctgtcatttctcagagagccctgcctctcaactgtgaaatgtaatgtaatg	2340
Db	2281	GAGAAATGTCTGCTCATTTTCAGAGAGCCCTGCTTCACACTGTGAGATGGTGGTGTG	2340
Qy	2341	gaaacccggtgtcaagttaatgtatgtgcagagagcagtaaaatgtccagccagtcagatagagt	2400
Db	2341	GAAAACCCGGTATAGGTATATGTATGTGTGGCAGGAGCACTCAAAATGCCAGCAATATGGGCT	2400
Qy	2401	gggtaccctgtgtaaacccaaccttaaaagctgaaagaagtcaccggcttaaacctctatactg	2460
Db	2401	GGGTACCTGGTGAACCCAACTTAAAGCTGAAGAGACGTCCGGGCTAAATCCTCATATG	2460
Qy	2461	aattgaaacccgtctctcccatctgtgtgtcttccctccagatgaatccaaacctcaacc	2520

OY	2587	cccttgacacttaccaaatcagtlcaacgltgtatcttccctatctctgagcccaataaagacc	2646
Db	124781	cccttgacacttaccaaatcagtlcaacgltgtatcttccctatctctgagcccaataaagacc	124722
OY	2647	cagatcagctgcaatggaggagaagaatcacccctctctgtggaggtgtggagaccatccct	2706
Db	124721	cagatcagctgcaatggaggagaagaatcacccctctctgtggaggtgtggagaccatccct	124662
OY	2707	gataccccctcctcactggaagctgtctcttcttcctcaataaattcttctccatccatct	2766
Db	124661	gcataccccctcctcactggaagctgtctcttcttcctcaataaattcttctccatccatccat	124602
OY	2767	caacct 2772	
Db	124601	caacct 124596	
RESULT	4		
ACOL6590			
LOCUS			
DEFINITION	ACOL6590	253217 bp	DNA linear HTG 20-APR-2001
ACCESSION	Homo sapiens chromosome 19 clone CTD-3220F14, WORKING DRAFT		
VERSION	ACOL6590		
KEYWORDS	ACOL6590 6	GI:13699590	
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	DOE Joint Genome Institute.		
REFERENCE	2 (bases 1 to 253217)		
AUTHORS	Unpublished		
TITLE	DOE Joint Genome Institute.		
JOURNAL	2 (bases 1 to 253217)		
REFERENCE	DOE Joint Genome Institute.		
AUTHORS	Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint		
TITLE	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
JOURNAL	On Apr 20, 2001 this sequence version replaced gi:7711568.		
COMMENT	-----Genome Center		
	Center: Joint Genome Institute		
	Center Code: JGI		
	Web site: http://www.jgi.doe.gov		

	Project Information		
	Center Project Name: 940643, BC905667		
	Center clone name: CITB-EL_3220F14		

	Summary Statistics		
	Consensus quality: 207648 bases at least Q40		
	Consensus quality: 225569 bases at least Q30		
	Consensus quality: 232428 bases at least Q20		
	Estimated insert size: 250510; agarose-1p estimation		
	Estimated insert size: 249417; sum-of-contrigs estimation		
	Quality coverage: 9.03 in Q20 bases; agarose-1p estimation		
	Quality coverage: 9.07 in Q20 bases; sum-of-contrigs estimation.		
	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 39 contrigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contrigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		
	1	1019: contrig of 1019 bp in length	
	*	1020	1119: gap of unknown length
	*	1120	2152: contrig of 1033 bp in length
	*	2153	2252: gap of unknown length
	*	2253	3701: contrig of 1449 bp in length
	*	3702	3801: gap of unknown length
	*	3802	5053: contrig of 1252 bp in length
	*	5054	5153: gap of unknown length

*	5154	6184:	contig of 1031 bp in length
*	6185	6284:	gap of unknown length
*	6285	7804:	contig of 1525 bp in length
*	7810	7909:	gap of unknown length
*	7910	8986:	contig of 1077 bp in length
*	8987	9086:	gap of unknown length
*	9087	10090:	contig of 1004 bp in length
*	10091	10190:	gap of unknown length
*	10191	11225:	contig of 1035 bp in length
*	11226	11325:	gap of unknown length
*	11326	12365:	contig of 1040 bp in length
*	12366	12465:	gap of unknown length
*	12466	14535:	contig of 2070 bp in length
*	14536	14635:	gap of unknown length
*	14636	16489:	contig of 1854 bp in length
*	16490	16589:	gap of unknown length
*	16590	19109:	contig of 2520 bp in length
*	19110	19209:	gap of unknown length
*	19210	21841:	contig of 2632 bp in length
*	21842	21941:	gap of unknown length
*	21942	23735:	contig of 1794 bp in length
*	23736	23835:	gap of unknown length
*	23836	25861:	contig of 2026 bp in length
*	25862	25961:	gap of unknown length
*	25962	29117:	contig of 3156 bp in length
*	29118	29217:	gap of unknown length
*	31908	32007:	contig of 2690 bp in length
*	32008	36364:	contig of 4357 bp in length
*	36365	36464:	gap of unknown length
*	36465	43309:	contig of 6845 bp in length
*	43310	43409:	gap of unknown length
*	43410	48185:	contig of 4776 bp in length
*	48186	48285:	gap of unknown length
*	48286	55626:	contig of 7341 bp in length
*	55627	55726:	gap of unknown length
*	55727	65781:	contig of 10055 bp in length
*	65782	65881:	gap of unknown length
*	65882	69673:	contig of 3792 bp in length
*	69674	69773:	gap of unknown length
*	69774	79389:	contig of 9616 bp in length
*	79390	79489:	gap of unknown length
*	79490	91483:	contig of 11964 bp in length
*	91484	91553:	gap of unknown length
*	91554	96671:	contig of 5118 bp in length
*	96672	96771:	gap of unknown length
*	96772	100544:	contig of 3773 bp in length
*	100545	100644:	gap of unknown length
*	100645	111423:	contig of 10779 bp in length
*	111424	111523:	gap of unknown length
*	111524	115278:	contig of 3755 bp in length
*	115279	115378:	gap of unknown length
*	115379	124011:	contig of 8633 bp in length
*	124012	124111:	gap of unknown length
*	124112	138475:	contig of 14364 bp in length
*	138476	138575:	gap of unknown length
*	138576	140613:	contig of 2038 bp in length
*	140614	140713:	gap of unknown length
*	140714	144949:	contig of 4236 bp in length
*	144950	145049:	gap of unknown length
*	145050	164352:	contig of 19303 bp in length
*	164353	164452:	gap of unknown length
*	164453	170540:	contig of 6088 bp in length
*	170541	170640:	gap of unknown length
*	170641	185046:	contig of 14406 bp in length
*	185047	185146:	gap of unknown length
*	185147	222209:	contig of 37063 bp in length
*	222210	222309:	gap of unknown length
*	222310	253217:	contig of 30908 bp in length
Location/Qualifiers			
1..253217			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			

```
1. .253217
/organism="Homo sapiens"
/db_xref="taxon:9606"
```


OY 2467 gaactgtctccatttggtgtgtcttcctcgaattgatcccaacccttaccatttt 2526
|||||
Db 76551 GAACCTGCTTCCCATTTGGTGTGCTTCTCCGATGATCCCAACCTTCACCTATTTT 76610
OY 2527 acgtatactgccccttccctaatttggttttacctgctgtgcccaccltttgatggtg 2586
Db 76611 ACCTATACCTGCTTCTTCTTATTTGTTTACATGCTGTGCCCCACCTTTTGAAGTGAG 76670
OY 2587 ccttggcattacttaaatcagtaacgtgtatctccctattcttgagccataaaagacc 2646
|||||
Db 76671 CTTTGCACTACTTACAAATCACTCAACGCTATTCCTCCCTATTCCTAGCCCATMAAGAC 76730
OY 2647 cgaactcagctcagtgaggaagaatacaacctgtctgtgaggttgaggaaccacct 2706
|||||
Db 76731 CAGACTCACTGCTCACTAGAGAGAGAAATCACCTGCTGTGGGGGTTTGGGACACTCCCT 76790
OY 2707 gcatccctctccactgagagctgtcttcttgctcaataaattcttcttaccatcct 2766
|||||
Db 76791 GCATCCCTCTCCACTGAGAGCTGTCTTCTTCTCAATMAATCTTTTCTACCATCCT 76850
OY 2767 caccct 2772
|||||
Db 76851 CACCT 76856

RESULT 5
AC093235 38173 bp DNA linear PRI 29-SEP-2001
LOCUS AC093235
DEFINITION Homo sapiens chromosome 19 clone LNLNR-267A4, complete sequence.
ACCESSION AC093235
VERSION AC093235.2 GI:15808549
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 38173)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 38173)
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 38173)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
COMMENT Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Sep 29, 2001 this sequence version replaced gi:15193369.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source
1. 38173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="LNLNR-267A4"
BASE COUNT 9495 a 8465 c 8118 g 12095 t
ORIGIN

Query Match 79.1%; Score 2193; DB 9; Length 38173;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2343; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 427 cagaatcgagccagaataactagctccctcctgctcctctgatttctccagtcagc 486

|||||
Db 28888 CAGATTCAGAGCAAAATCAACTTAGTCCCTCTGCTCTGTATTTCTCCAGTCAGC 28947
OY 487 aagcctcagccaaatgctgtgctgagtcattctctcagctgttctcaagttatgg 546
Db 28948 AAGCTCTAGCCCAACATGTGTGCTGAGTCACTCTCTCAGCTGTTTCAATTTATGGG 29007
OY 547 caggaaatccctccacctgggaaacacatccagaagatcagaagaacagcagagtc 606
|||||
Db 29008 CAGGAATCTCTCCACTGGGAAACATATCCAGAAATCAGAAACACGACGAGATC 29067
OY 607 cacttgcctttagtggcaagagaatgattcaagaggaggaagacccctcagctcctgt 666
Db 29068 CATCTGCTTTAGTGGCAAGAGAGAAATGATTCAGAGGAGAGAACTCCAGACTCTGT 29127
OY 667 ttggagagtgtaagcaaaaatggtcacttcaagagccttccagccacctgaagaacac 726
Db 29128 TTGGAGAGTAAAGCAAAATGGCACTTCAAAAGCACTTCCAGCCCTGAAAGAACAC 29187
OY 727 agccagcagctccagaggagaacacacagtggtgatataggtccagccctgaagga 786
Db 29188 AGCCAGCAGCTCCAGAGAGACAAACACAGTGTGATATAGGGTCCAGCCTGAAACGGA 29247
OY 787 gggcagatctagaggaacacagacaagatltgcatgtttagaagtcacagatttgag 846
Db 29248 GGGCAGATCTAGAGGAACAGACAAAGTATTCATGTGTTAGAGTCTCAGGATTTGGAG 29307
OY 847 aatacaatatagaagatttgggtccaggtcttatacagaggtcacaacctctcctaccc 906
Db 29308 AAATCAAAATATGAAAGATTTGGGCCAGCTTTATCAAGAGAGCAAACTCTTACCTCC 29367
OY 907 aagaagacaacactggggagacaccttaccatgtacactgaagtggggagagacgttgca 966
Db 29368 AGAAGACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGAGACACTTTGGCA 29427
OY 967 gtagtcaagctcctcaaaaaacccaaggaacaccttgggggaaagccttatgtgtca 1026
Db 29428 GTATGTCAGTCTCTCATCAAAAACCCAGAGACACACTGTGGGGAAGCTTTATGTGTGA 29487
OY 1027 gggaaatgtgggagaggtttagtgaagttcaaacctgaatcaaatcagaagaacact 1086
Db 29488 GGGAAATGTGGGAGAGCTTTAGCTGAGATCAAACTGTATCAACATTCAGAGACACACT 29547
OY 1087 caggaggaagaacctatgtgtgcaagagatltgcaagaggttcttacttggaaagtcgaacc 1146
Db 29548 CAGGGGAGAACTTATGTGTCAAGAGATTGGACGAGCTTTACTTGGAAAGTGAAC 29607
OY 1147 tctttacacatcagcgagacacatcagggtccttaagtgtgtgcaaggaattgtggc 1206
Db 29608 TCTTTACACATCAGCGGACACACTCAGGGCTCAAGCTTATGTGTGCAAGGAATGTGGGC 29667
OY 1207 aagactttagcctgaagttcaaacctcatccacagaagggtcggaacttgggggaagc 1266
Db 29668 AGAGCTTTAGCTTGAAGTAACTCTATTACCCACAGGGGCGCACATGGGGAGAAAC 29727
OY 1267 ctatgttgcaggaagtgtggcgtgtgcttgcgcaagatccacacacctgtgtcagacaca 1326
Db 29728 CTTATGTTGACGAGGAATGTGGCGGTTCGCGCAGATTCACACCTGTGTGACAGACA 29787
OY 1327 agaggaacacatcaggaagaagaccttaccatttgcagggagtgtgagcaaggtttagcc 1386
Db 29788 AGAGGACATTCAGAGAGAGAGAGCCCTTACATTGTGCGGAGGTGTGAGCAAGCCTTAGCC 29847
OY 1387 aagaagtcacacctatcagacacttaagaagacacacagagagaagaccttatgtatga 1446
Db 29848 AGAAGTCACACTCATCAACACCTTAAGGACACACACAGAGAGAAAGCTTTATGTATGCA 29907
OY 1447 cagaatgtgggtcactttagcttggaatcaaacctcacaacccagagagacact 1506
Db 29908 CAGAAATGTGGCATACCTTAGCTGGAATCAAACTCAAAACACACCGAGAGACACT 29967
OY 1507 cagggttaaaccttatgtctgctgaggtgtcggtgaggtgtttagcctggaagtcacac 1566
|||||

Db	29968	CAGGGTTAAACCTTATGTCTGCTGCTGAGATGGGGGACAGTGTCTTAACTGAAAGTCAAAACC	30021
Qy	1567	tttaacaaacacagagtgatcacacacggyggaagaccatttgaatgacgaatgtgagc	1626
Db	30028	TTTAAACAACACAGAGGTGACACACAGGGGAGAGCACTTTGTATGTAGTGGAGTGGGC	30087
Qy	1627	gaggtcttaaccccggaataacaccgtatcagcaccagaagagacacactaagyygaagac	1686
Db	30088	GGGGCTTTTACCCGGAAATACCTGTAGACGCACACAGAGGACACACTAGGGGAGAAAC	30147
Qy	1687	caattgtaatgtcgtagatgtgagcagagctttaaataagctcaaccccatcttaaac	1746
Db	30148	CATTGTATGTGCTGAGTGTGACAGCAGGCTTTAAATGATATAGTCCACTTCATTTCACAC	30207
Qy	1747	agaagacacatttcagggagaaaaccccttataatgacgaagatgtgccaagaagcttcgcgc	1806
Db	30208	ACAGGACACATTTCAGGGGAAAGCCTTTATGTGACGGAGATGTGGCAGAAAGTTTGGC	30267
Qy	1807	agaagacctaacctgttttaagcacaagaaggacacactcaagtgccctgtgtgcaaggaat	1866
Db	30268	ACAAGCCTTAACTGTTTATAGGCACAAGAGGGCACACTCAAGTCCCTTGTGTGACAGGAGT	30327
Qy	1867	gttgggcaagagcctttgtgtgtaagttaactcttaataaacaagaagagagacacgaaggg	1926
Db	30328	GTGGGCAAGGCTTGTGTGCTAAGTTAACTCTATTAAACACCAAGAGACACAGCAGGGG	30387
Qy	1927	ggaagccctcatgtgtgcaaggaatgtggycaagagctttagccggaagtacaccccatla	1986
Db	30388	GGAAGCCTATGTGTGACAGGAGATGTGGGCAAGGCTTATAGCCGGCAGTCAACCTCATTA	30447
Qy	1987	gaaccacaagagacacatttcagaagagaacccctataattgtcagaagaatgttgacggygct	2046
Db	30448	GACACCAGAGGACACATTCAGAGAGAAAGCCTTTATTTTGCAGAAAGTGTGACGGGCT	30507
Qy	2047	ttaagtcgaagctccaaccttatcaacatcagaagacacaccctcagaatagaaccttaly	2106
Db	30508	TTAGTGTGGAAGTCAACCTTATCAACAACATCAGAGACACACACAGAGATGAAGAACTTTATG	30567
Qy	2107	tgataaggaatgtgtgtacagcccttaagccaagagtcatacttcaacgaacacagaaga	2166
Db	30568	TGTATAGGGAATGTGTGTACAGGCTTTAGCCAGAGATCATCTTCAATCAACACACAGAGA	30627
Qy	2167	cacacacagtgctgtgtgcttlttcagcaatgtctagataccaagaatgtgagacatctgtg	2226
Db	30628	CACACACAGTGTGTGGCTTTTTCAGCAATTCCTGTATACCAAAAGTGAGACATTCGTG	30687
Qy	2227	tgtgttatgtcatgtgagctcgtacatgtgttaagctgttatactccactcgaagaagaat	2286
Db	30688	TGTGATTTATGTGATGAGCTGATCTATCTGTAAAGCTTTATCTCCATCCACTGAAAGAGAT	30747
Qy	2287	tgctgtgcatcttcaagagacccgtccctctccactgtgaatgtgtgtgtgtgtgaaac	2346
Db	30748	TGCTGTCTCATTTTTCAGGAGCCCTCCCTCTCACTGTGATGTGGTGTGTGAAMCC	30807
Qy	2347	cggctcagtaabgatagtgtgcagaagagcaatcaaatgtcccagcagataaggtgtggtac	2406
Db	30808	CGGTGAGTAAATGATAGTGGCGAGGAGCAAGTCAATATGCCACAGCAGATAGGGGTGTGAC	30867
Qy	2407	ctgtgtgaaacccaacttaaaagctgtaaagaacgttcccggtcaaatctccatactgtaatgtga	2466
Db	30868	CTGGTGAACACCCAACTTAAAGCTTGAACAACATCCCGGCTTAATTCCTATGCTGAATTGA	30927
Qy	2467	gaacctgtcctcccaattgtgtgtgtcttcctcccgatgatcccaaacccctcaaccatttt	2526
Db	30928	GAACTGTCTTCCCATTTTGGTGTGCTTTCCTCCGATTATCCCAACCTTCACTATTTTT	30987
Qy	2527	acgtataccctgcctcttccataatgtgtttttaaacaatgtgtgccaacctttgagtgtg	2586
Db	30988	ACGTATACCTGCGCCCTTCCCTAATTGTTTAAACATGCTGTGTGCCACCTTTTGTAGTGTG	31047
Qy	2587	ccttgcataacttacaatcaagtcagtaacgtgtatctcccatlcttgagccataaagaacc	2646
Db	31048	CCTTTGTCTACTTCAAAATCAGTCAAACTGTATTTCCCTATTCTTGAGGCCCATAAAGAACCC	31107

QY	2647	cgagctcagctcgaatgagagagagaataacacccctgctgtgagaggttgggagaccatccct	2706
Db	31108		31167
QY	2707	gcattccctccacactgagagctgtctcttctgctaataaattcttctcaccatcct	2766
Db	31168		31227
QY	2767	caacct 2772	
Db	31228		
		CACCTT 31233	
RESULT	6		
LOCUS	BC004513	1998 bp mRNA linear PRI 12-JUL-2001	
DEFINITION	Homo sapiens, clone IMAGE:3928207, mRNA, partial cds.		
ACCESSION	BC004513		
VERSION	BC004513.1	GI:13325426	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1998)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submissions		
	Submitted (12-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTD/DRP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadane@systemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahney, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan		
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 14 Row: b Column: 10 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.</p>		
SOURCE	<p>Location/Qualifiers</p> <p>1..1998</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:3928207"</p> <p>/tissue_type="Ovary, adenocarcinoma"</p> <p>/clone_lib="NIH_MGC_9"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOTB7"</p> <p><1..1280</p> <p>/codon_start=3</p> <p>/product="unknown (protein for IMAGE:3928207)"</p> <p>/protein_id="AAH04513.1"</p> <p>/db_xref="GI:13325427"</p> <p>/translation="HEVSVSGFGEIKYEEFPGFIKESNLISLQKTOTGEMPYWTEK/GDSGNSVLINKNPRTHSGGKPYVRCREGGRFTKSNLITHQRTSHGKPYVCDGGRGFMTWSNLFTHQRTHSLKPYVCEKCGSFSLKSNLITHQRAHNGKPYVRCGGRGFROSHLIVRHKRTHSGKPYVRCGCGSOKSLIRHRTTGKRPVCTGCGHFRFSKSNLKTQRTSHSGKPYVLECGGCFSLKSNLKHQSHNGKRPVCTGCGHFRFSKSLITHQRTSHSGKPYVCAECGRFNDSTLSLHQRTHSGKPYVRCGGRFROKPNLFRHRAHSGAIVRCGGRFCAPKLTLLKHQRAHNGKPYVRCGGRGFSROSLIRHQRTHSGKPYVRCGRGFSRKSNIIRQRTSHSG"</p>		
CDS	<p>575 a 458 c 510 q 455 t</p>		
BASE COUNT			

ORIGIN

Query Match 63.9%; Score 1771; DB 9; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 826 tagaagctccagagatttggaagaatcaataatgaagagcttggccagagcttatcaag 885
DB 10 TAGAAGCTCCAGGATTTGGAGAAATCAATATGAGAGTTTGGCCAGGCTTATCAAG 69
QY 886 agtcaaacctccttagcctccagaagacacaaactlgyggagacacttacaatgtaactg 945
DB 70 AGTCAAACTCCTTACCTCCAGAAACACAACTGGGAGACACTTACATGTACACTG 129
QY 946 agtgggaaagacagcttggcagatgatcgactcctcaaaaaaaccagaagacactctg 1005
DB 130 AGTGGGAGACAGCTTGGCAGTATGTCACTCCTCATCAAAAAACCAAGGACACTCTG 189
QY 1006 gggaaagacctatgtgtcagggaaatgtggtcgaagcttcaatgtaagtcacactga 1065
DB 190 GGGAAAGCCTTATGTGTCTCAGGAGATGTGGGAGAGCTTACGTGGAAGTCAAACTGA 249
QY 1066 tcaacacatcagaggacacactcagggagaaacttatgtgtcagaagatgtgtgacag 1125
DB 250 TCACACATCAGAGACACACTCAGGGAGAAACCTTATGTGTCAAGGATTTGTGACAG 309
QY 1126 gcttaacttggaagtcgaacctcttaacacatcagagacacactcagagctcaaacctt 1185
DB 310 GCTTACTTGAAGTGAACCTTTTACATCAGCGACACTCAGGCTTCAAACTCA 369
QY 1186 atgtgtcagaagaaatgtgtgacagagctttagcctgaagtcacaaactcattaccacaga 1245
DB 370 ATGTGTGCAAGGAATGTGGGACAGCTTATGCTTGAAGTCAAACTCATTAACACACAG 429
QY 1246 gggcgacacctggggagaaagccttatgtgtcagggaaatgtgtgctgtgtcgcacg 1305
DB 430 GGGCGACACTGGGGAAGCCTTATGTTGCAAGGAATGTGGGCTGTGGCTTTCGCGACG 489
QY 1306 attcacactgtgtcagaacacaaagagacacatcagagagaaagccttacaattgtcagg 1365
DB 490 ATTCAACCTGCTGCAGACACACAGAGACACATTCAGAGAGAAAGCCTTATTCAGAGG 549
QY 1366 agtgtgagcaagctttagccagaagtcacacactcattcagacacttaagagacacacag 1425
DB 550 AGTGTAGCAAGCTTTAGCCAGAGATGCACACTCATCAAGACACTTAAGAGACACACAG 609
QY 1426 gagagaagccttatgtatgcagaagaatgtgtgctcactttagctgtgaaataaacctca 1485
DB 610 GAGAGAAGCCTTATGTATGCACAGAAATGTGGGCTCACTTACCTGGAATCAAACTCA 669
QY 1486 aaacacacagagagacacactcaggggtttaaacttatgtgtcctcggagttggcgact 1545
DB 670 AAACACACAGAGACACACTCAGGGGTTTAACTTATGTCTCTGAGACTGCGGGGAGT 729
QY 1546 gctttagcctgaagtcacaaactcaacacacagagatcacacacagggggagaaagccat 1605
DB 730 GCTTATGCGTAAGTCAAACTTAAACACACAGAGGTACACAGGGGAGAAAGCCAT 789
QY 1606 tgtatgtacagagatgtgtggaagcttaccggaaatacaacctgtatcagcaccaga 1665
DB 790 TGTATGTACGAGTGTGGGCGAGGCTTAACTCCGGAATCAACCCGTATCACCACACAGA 849
QY 1666 ggaacacacagggagaaagcaattgtatgtgtcagttgtgagagagctttaaagata 1725
DB 850 GGAACACACTCAGGGGAGAAAGCCATTGTATGTCTGAGTGTGACGAGGCTTAAATGATA 909
QY 1726 agtcaaccctcatttcaacacagagagacacatcagggagaaagcctttagtgcagag 1785
DB 910 AGTCAACCTCATTTTCAACACAGAGACATTCAGGGGAAAGCCTTTATGTGACAGG 969
QY 1786 agtgtggcagaaggttcggcagaagcctaactgtttaggcacaagaggcactcag 1845
|||||

DB 970 AGTGTGGCAGAAGGTTTCGGCAGAAAGCCTTAACCTGTTTAGGCAAGAGGACACTCAG 1029
QY 1846 gtgccttgtgtcagggagatgtggcaaggtcttgtgtctaagttaactcattcaaac 1905
DB 1030 GTGCCTTGTGTGACGGAGATGTGGCAAGGCTTTGTGTCTAAGTTAACTTCATTAAC 1089
QY 1906 accagagacacacagcagggggagagcctcattgtgtcagggagatgtgtgcaagcttca 1965
DB 1090 ACCAGAGACACACGACGGGGGAAAGCCTCATGTGTGCAAGGAGTGTGGCAAGGCTTAA 1149
QY 1966 gccggcagtcacacccatcattagacaccagagacacatcagaagagaagccttattt 2025
DB 1150 GCCGCGAGTCACACCTCATTTAGACACAGAGACACATTCAGAGAGAAAGCCTTATATTT 1209
QY 2026 ggaagaagtgtgagcggccttaagtgcgaagtccaacttccagacacagaagagac 2085
DB 1210 GCAGAAAGTGTGACGGGCTTTAGTCGAAAGTCAACCTTATCAACATTCAGAGACAC 1269
QY 2086 actcagatagaaactttagttagaggaatgtgtacagcctttagccagagatca 2145
DB 1270 ACTCAGATAGAAACTTATGTGTATAGGGAATGTGTACAGCCTTTAGCCAGAGACTCAT 1329
QY 2146 acttcaacagacacacagagacacacacagtgctgtgtctttagccattgtcagata 2205
DB 1330 ACTTCATCAGACACCCAGAGACACACAGTGTGTGCTTTTTCAGCCATTGCTAGATA 1389
QY 2206 ccaaaatgagacatctgtgtgtatgtacatgagacgtgactgtgttagactgtatc 2265
DB 1390 CCMAATGGAGACATTTCTGTGTGTATGATGATGAGACTGTCTGTAAAGACTTATTC 1449
QY 2266 tccatccacctgaaagaaatgtgtgtcatttcaaggaagcctgtccctccacatgt 2325
DB 1450 TCCATCCACCTGAAGAGAAATTCCTGCTCATTTTCAGAGCCCTCCCTCCTCACTGT 1509
QY 2326 gtagtgtgtgtgtgtgaaacccggtcaggttaagtatgagtgagcagggaggaatgaatgc 2385
DB 1510 GGAATGTGGGTTGT-6AAACCCGGTCAAGTAATGATGTGTGCGAGAGCAAGTCAAAATGCC 1568
QY 2386 caagcagatagagggtgtgtaccgtgtgtgaacccaactttaaagcttaaagacagttcccg 2445
DB 1569 CAGCGAGATAGGGGTGTGTGATGCTGTGTGAACCCACACTTAAAGCTGAAGACATGCCGG 1628
QY 2446 taaatcccatcactgaattgagaacctgtctcccatgtgtgtgtgtcctccatgtga 2505
DB 1629 TAAATCCTCATACTGAATTGAGAACTGTCTCCATTTGTGTGTGCTTCCCTCCGATTGA 1688
QY 2506 tcccaaaccttcaactatlttagtataactgtcccttccctaaattgtttttaaactgtc 2565
DB 1689 TCCCAACCTTCACTATTTTATGATATACCTGCTCCCTTCTTAATGTGTTTAACTGTCT 1748
QY 2566 gtgccacacctttagatgtgtgtccttgcataactcaaatcaatcaacgtgtatctccct 2625
DB 1749 GTGCCACCTTTTGAATGTGTGCTTGTGCATACTTACAATCAATCAAGATGTTCCCT 1808
QY 2626 attctgaagccataaaagaccagactcagctgtcagtgagagagaatacaacctgtctg 2685
DB 1809 ATTCTAGCCCATTAATAAGACCCAGACTCAGCTCAGAGAGAGAAATCAACCTGCTGT 1868
QY 2686 ggaagtgtggagacactccctgtcatccctctcactcgaagatgtgtcttgtctaata 2745
DB 1869 GGGGGTGGGACACCTCCCTGATCCCTTCACACTGAGAGCTGTCTTTTGTCTCAATA 1928
QY 2746 aaattccttctaccacactca 2768
DB 1929 AAATTCCTTCTACCATCCTCA 1951
|||||

RESULT 7
HUMHR1 HUMHR1 672 bp DNA linear PRI 27-APR-1993
LOCUS Human Kruppel related gene, exon X, clone pKRR18.
DEFINITION M20675.M19429
ACCESSION M20675.1 GI:184108
VERSION

KEYWORDS Krueppel-related protein.
SOURCE Human DNA, clone pKRIIRS.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 672)
AUTHORS Ruppert, J.M., Kinsler, K.W., Wong, A.J., Bigner, S.H., Kao, F.-T.,
Law, M.L., Senanez, H.N., O'Brien, S.J. and Vogelstein, B.
TITLE The Gli-Kruppel family of human genes
JOURNAL Mol. Cell. Biol. 8, 3104-3113 (1988)
MEDLINE 89096896
COMMENT computer-readable sequence for [1] kindly provided by J.M.Ruppert,
20-JUN-1988.
FEATURES
source location/Qualifiers
1..672
/organism="Homo sapiens"
/db_xref="taxon:9606"
exon <1..>672
/note="exon X"
BASE COUNT 195 a 155 c 184 g 138 t
ORIGIN Chromosome 19.
Query Match 24.28; Score 672; DB 9; Length 672;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 942 actgaatgaggagagacattgagcatatgcatcctcatcaaaacccaagagacac 1001
|||||
Db 1 actgagtgaggagacacgtttggcagatgctcactccatcaaaacccaagagacac 60
QY 1002 tctgggggaaagccttatgtgtgcaaggaaatgtggcgaggtcttaagtgaatcaac 1061
|||||
Db 61 tctgggggaaagccttatgtgtgcaaggaaatgtggcgaggtcttaagtgaatcaac 120
QY 1062 ctgatcacatcagagagacactcaggggagaacacctatgtgtgcaaggatgtgga 1121
|||||
Db 121 ctgattcagacatcagagagacactcaggggagaacacctatgtgtgcaaggatgtgga 180
QY 1122 cgaaggttacttggaagtgcaacctctcttaacatcagcgagagacacacacacac 1181
|||||
Db 181 cgaaggttacttggaagtgcaacctctcttaacatcagcgagagacacacacacacac 240
QY 1182 ccttatgtgtgcaaggaaatgtggcgaggtcttaagtgaatcaaaacctatcaaac 1241
|||||
Db 241 ccttatgtgtgcaaggaaatgtggcgaggtcttaagtgaatcaaaacctatcaaac 300
QY 1242 cagagggcgacacactggtggggaagccttatgtgtgcaaggaaatgtggcggtcttcgc 1301
|||||
Db 301 cagagggcgacacactggtggggaagccttatgtgtgcaaggaaatgtggcggtcttcgc 360
QY 1302 caagatcacacactgtgcaacacacacacacacacacacacacacacacacacacac 1361
|||||
Db 361 caagatcacacactgtgcaacacacacacacacacacacacacacacacacacacac 420
QY 1362 agggagtgtagaagccttaagcagaagtacacacacacacacacacacacacacac 1421
|||||
Db 421 agggagtgtagaagccttaagcagaagtacacacacacacacacacacacacacac 480
QY 1422 aacggagagaagccttatgtgtgcaaggaaatgtggcggtcttcacacacacacac 1481
|||||
Db 481 aacggagagaagccttatgtgtgcaaggaaatgtggcggtcttcacacacacacacac 540
QY 1482 ctcaaaac 1541
|||||
Db 541 ctcaaaac 600
QY 1542 caatgtcttagccttgaagtcacacacacacacacacacacacacacacacacacacac 1601
|||||
Db 601 caatgtcttagccttgaagtcacacacacacacacacacacacacacacacacacacac 660
QY 1602 ccattgtatgt 1613
|||||

Db 661 CCATTGTATGT 672
RESULT 8
LOCUS AX067361 678 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 65 from patent WO0078960.
ACCESSION AX067361
VERSION AX067361.1 GI:12544985
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS Vignin, J. and Mitcham, J.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0078960-A 65 28-DEC-2000;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..678
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..678
/note="n = A,T,C or G"
BASE COUNT 165 a 167 c 160 g 182 t 4 others
ORIGIN
Query Match 20.6%; Score 571; DB 6; Length 678;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2078 gaggacacactcagagatagaacttatgtgtataggaagtgtgtacagccttagaca 2137
|||||
Db 9 gaggacacactcagagatagaacttatgtgtataggaagtgtgtacagccttagaca 68
QY 2138 ggaagtcacttcatcagacacacagagagacacacacacacacacacacacacac 2197
|||||
Db 69 ggaagtcacttcatcagacacacacagagagacacacacacacacacacacacacac 128
QY 2198 gctagataccagaagtggagacatctgtgtgtatattgcatgagactgtactgtga 2257
|||||
Db 129 gctagataccagaagtggagacatctgtgtgtatattgcatgagactgtactgtga 188
QY 2258 ctgtatctccatccacactgaagagagaaatgtgtgtgtgtgtgtgtgtgtgtgtgt 2317
|||||
Db 189 ctgtatctccatccacactgaagagagaaatgtgtgtgtgtgtgtgtgtgtgtgtgt 248
QY 2318 ctcaactgtgatgt 2377
|||||
Db 249 ctcaactgtgatgt 308
QY 2378 caaatgcccagagacagataggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2437
|||||
Db 309 caaatgcccagagacagataggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 368
QY 2438 gtcccggtcaataccctcatatcagatgtgaacactgtctcccaattgtgtgtgtgt 2497
|||||
Db 369 gtcccggtcaataccctcatatcagatgtgaacactgtctcccaattgtgtgtgtgt 428
QY 2498 ccgattgaaccac 2557
|||||
Db 429 ccgattgaaccac 488
QY 2558 acaactgtgtccac 2617
|||||
Db 489 acaactgtgtccac 548
QY 2618 atccacctatctgagccacataaagaccca 2648
|||||
Db 549 attcccttatctgagccacataaagaccca 579

[illegible]

	*	52260	55732:	gap of unknown length
	*	55733	57755:	gap of unknown length
	*	57757	61389:	gap of unknown length
	*	61330	65398:	gap of unknown length
	*	65399	69740:	gap of unknown length
	*	69741	74577:	gap of unknown length
	*	74578	80165:	gap of unknown length
	*	80166	86055:	gap of unknown length
	*	86056	89215:	gap of unknown length
	*	89216	94479:	gap of unknown length
	*	94480	103136:	gap of unknown length
	*	103137	113254:	gap of unknown length
	*	113255	130351:	gap of unknown length
FEATURES	*			
source			Location/Qualifiers	
			1. .130351	
			/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
			/chromosome="3p"	
			/clone="RP11-420a23"	
BASE COUNT	38785	a	26050	c 26300 g 39176 t 40 others
ORIGIN				

```
Query Match      1.3%; Score 37; DB 2; Length 130351;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 2462 atbagaacctgtctcccatattggtgtcttcctc 2498
|||||
Db 69153 ATTGAGACCTGTCTTCCCAATTGGTGTGCTTCCCTC 69189

RESULT	10
AC096898	
LOCUS	
DEFINITION	AC096898 Homo sapiens chromosome 4 clone RP11-618K19, WORKING DRAFT
ACCESSION	AC096898
VERSION	AC096898.5 GI:18139542
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 161625)
JOURNAL	Waterston, R.H.
REFERENCE	The sequence of Homo sapiens clone
AUTHORS	unpublished
TITLE	2 (bases 1 to 161625)
JOURNAL	Waterston, R.H.
COMMENT	Direct Submission Submitted (01-OCT-2001) Genome Sequencing Center, Washington University school of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Jan 12, 2002 this sequence version replaced gi:17921273.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: H_NH0618K19
 ----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-terminator; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 160310 bases at least Q40
 Consensus quality: 160620 bases at least Q30
 Consensus quality: 160738 bases at least Q20
 Insert size: 161425; sum-of-contigs
 Quality coverage: 7.76 in Q20 bases; agarose-fp
 Quality coverage: 8.11 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 3023: contig of 3023 bp in length
 * 3024 3123: gap of unknown length
 * 3124 115853: contig of 112730 bp in length
 * 115854 115953: gap of unknown length
 * 115954 161625: contig of 45672 bp in length.
 Location/Qualifiers
 1. 161625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-618K19"
 1. 3023
 /note="assembly_name:Contig24"
 clone_end:SP6
 vector_side:left"
 3124. 115853
 /note="assembly_name:Contig26"
 clone_end:T7
 vector_side:right"
 115954. 161625
 /note="assembly_name:Contig25"
 BASE COUNT 47603 a 31872 c 31765 g 50185 t 200 others
 ORIGIN
 Query Match 1.3%; Score 37; DB 2; Length 161625;
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2462 atcgagaacctgtccattcgtggtgtccttcctc 2498
 ||||||||||||||||||||||||||||||||||||||||
 DB 110301 ATTGAGAACCTGCTTCCATTGCTGCTTCCCTC 110337
 RESULT 11
 AC034197 164264 bp DNA 11near HTG 29-MAY-2000
 LOCUS Homo sapiens chromosome 3 clone RP11-72801 map 3p, WORKING DRAFT
 DEFINITION
 SEQUENCE, 14 unordered pieces.
 AC034197.3 GI:8101273
 AC034197 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 164264)

AUTHORS
 Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
 Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
 Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
 Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y.,
 Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
 Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
 Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
 Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
 Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
 Zhang,Z., Zhu,B., Yu,J. and Yang,H.
 Chromosome 3p genomic sequence
 Unpublished
 2 (bases 1 to 164264)
 Zeng,Y., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
 Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
 Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
 Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
 Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
 and Yang,H.
 Direct Submission
 Submitted (05-APR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing
 100101, P.R.China
 On May 29, 2000 this sequence version replaced g1:7644466.
 -----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn
 ----- Project Information
 Center project name:1% project
 Center clone name: RP11-72801
 ----- Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 5% of reads
 Chemistry: Dye-terminator Big Dye; 45% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 15901 bases at least Q40
 Consensus quality: 163151 bases at least Q30
 Consensus quality: 165125 bases at least Q20
 Insert size: 158755; sum-of-contigs
 Quality coverage: 4.57x in Q20 bases;sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2043: contig of 2043 bp in length
 * 2044 2143: gap of unknown length
 * 2144 4407: contig of 2264 bp in length
 * 4408 4507: gap of unknown length
 * 4508 5590: contig of 1083 bp in length
 * 5591 5691: gap of unknown length
 * 5691 8894: contig of 3204 bp in length
 * 8895 8995: gap of unknown length
 * 8995 12938: contig of 3944 bp in length
 * 12939 13038: gap of unknown length
 * 13039 19151: contig of 6113 bp in length
 * 19152 19251: gap of unknown length
 * 19252 24784: contig of 5533 bp in length
 * 24785 24885: gap of unknown length
 * 24885 32377: contig of 7493 bp in length
 * 32378 32478: gap of unknown length
 * 32478 41545: contig of 9068 bp in length
 * 41546 41645: gap of unknown length
 * 41646 50878: contig of 9233 bp in length
 * 50879 50978: gap of unknown length
 * 50979 62812: contig of 11834 bp in length

```

*      *      *      *      *
*      62813    62912: gap of unknown length
*      62913    88788: contig of 25876 bp in length
*      88789    88888: gap of unknown length
*      88889    115986: contig of 27098 bp in length
*      115987    116087: gap of unknown length
*      116087    164264: contig of 48178 bp in length.
FEATURES
source
1..164264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-72801"
1..2043
/note="assembly_name:Contig4"
2144..4407
/note="assembly_name:Contig5"
4508..5590
/note="assembly_name:Contig6
clone_end:77
vector_side:right"
5691..8894
/note="assembly_name:Contig7"
8995..12938
/note="assembly_name:Contig8"
13039..19151
/note="assembly_name:Contig9"
19252..24784
/note="assembly_name:Contig10"
24885..32377
/note="assembly_name:Contig11"
32478..41545
/note="assembly_name:Contig12"
41646..50878
/note="assembly_name:Contig13"
50979..62812
/note="assembly_name:Contig14"
62913..88788
/note="assembly_name:Contig15"
88889..115986
/note="assembly_name:Contig16"
116087..164264
/note="assembly_name:Contig17"
BASE COUNT 46659 a 32302 c 32891 g 51080 t 1332 others
ORIGIN
Query Match 1.3%; Score 37; DB 2; Length 164264;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2462 attgagaacctgtctccattgtgtgtgtcttc 2498
|||||
Db 55166 ATTGAGAACCTGCTTCCTCCATTGCTGCTTCTTC 55202

RESULT 12
AC069265 165379 bp DNA 11linear HTG 29-MAY-2000
LOCUS Homo sapiens chromosome 3 clone RP11-208K18 map 3p, WORKING DRAFT
DEFINITION
SEQUENCE 7 unordered pieces.
AC069265
AC069265.2 GI:81011151
VERSION
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 165379)
Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Lu,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
Zhang,Z., Zhu,B., Yu,J. and Yang,H.
Unpublished
Chromosome 3p genomic sequence
2 (bases 1 to 165379)
Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
and Yang,H.
Direct Submission
Submitted (23-MAY-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On May 29, 2000 this sequence version replaced g1:8039668.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hg.citp.ac.cn
http://www.genomics.org.cn
Contact:hg@citp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-208K18
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 55% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 164602 bases at least Q40
Consensus quality: 165885 bases at least Q30
Consensus quality: 166968 bases at least Q20
Insert size: 162389; sum-of-coverage
Quality coverage: 6.82x in Q20 bases;sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1875 1874: contig of 1874 bp in length
1975 1974: gap of unknown length
1975 8441: contig of 6467 bp in length
8442 8541: gap of unknown length
8542 20983: contig of 12442 bp in length
20984 21083: gap of unknown length
21084 38280: contig of 17207 bp in length
38291 38390: gap of unknown length
38391 56761: contig of 18370 bp in length
56761 56860: gap of unknown length
56861 93526: contig of 36566 bp in length
93527 93627: gap of unknown length
93627 165379: contig of 71753 bp in length.
FEATURES
source
1..165379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-208K18"
1..1874
/note="assembly_name:Contig3"
1975..8441
/note="assembly_name:Contig4"
8542..20983

```

```
misc_feature      /note="assembly_name:Contig5"
                  21084..38290
                  /note="assembly_name:Contig6"
misc_feature      38391..56760
                  /note="assembly_name:Contig7"
misc_feature      56861..93526
                  /note="assembly_name:Contig8"
misc_feature      93627..165379
                  /note="assembly_name:Contig9"
BASE COUNT      51415 a 32065 c 32897 g 48394 t      608 others
ORIGIN

Query Match      1.3%; Score 37; DB 2; Length 165379;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2462 atgagacactgtctccattgtgtgtcttcctc 2498
      |||
Db 83850 ATTGAGAACTGCTTCCATTGCTGTCTTCCTC 83886

RESULT 13
AC026219      184375 bp      DNA      linear      HTG 29-MAY-2000
LOCUS      Homo sapiens chromosome 3 clone RP11-815J19 map 3p, WORKING DRAFT
DEFINITION      AC026219
ACCESSION      AC026219
VERSION      AC026219.2 GI:8101287
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 184375)
  Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
  Dong,W., Fan,H., Feng,X., Guan,Q., Guo,X., Guo,D., He,L., Hu,S.,
  Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
  Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
  Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
  Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
  Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
  Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
  Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
  Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  Chromosome 3p genomic sequence
  Unpublished
2 (bases 1 to 184375)
  Kang,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
  Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
  Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
  Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
  Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
  and Yang,H.
  Direct Submission
  Submitted (21-MAR-2000) Human Genomic Center, Institute of
  Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
  100101, P. R.China
  On May 29, 2000 this sequence version replaced gi:7272036.
  -----Genome Center
  Center:Beijing Center
  Center code:Beijing
  Website:http://hgsc.jgtp.ac.cn
  http://www.genomics.org.cn
  -----Project Information
  Project name:1% project
  Center clone name: RP11-815J19
  -----Summary Statistics
  Sequencing vector: pUC18: 100% of reads
  Chemistry: Dye-terminator: ET 55% of reads
  Chemistry: Dye-terminator Big Dye: 45% of reads
  Assembly program: Phrap: version 0.990329
```

```
Consensus quality: 146985 bases at least Q40
Consensus quality: 167323 bases at least Q30
Consensus quality: 182494 bases at least Q20
Insert size: 122192, sum-of-coverage
Quality coverage: 3.83x In Q20 bases;sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1389: contig of 1389 bp in length
1390 1489: gap of unknown length
1490 3602: contig of 2113 bp in length
3603 3702: gap of unknown length
3703 6147: contig of 2445 bp in length
6148 6247: gap of unknown length
6248 7841: contig of 1594 bp in length
7842 7941: gap of unknown length
7942 9569: contig of 1628 bp in length
9570 9669: gap of unknown length
9670 11970: contig of 2301 bp in length
11971 12070: gap of unknown length
12071 13209: contig of 1139 bp in length
13210 13309: gap of unknown length
13310 14906: contig of 1597 bp in length
14907 15006: gap of unknown length
15007 16915: contig of 1909 bp in length
16916 17015: gap of unknown length
17016 18655: contig of 1640 bp in length
18656 18755: gap of unknown length
18756 20974: contig of 2219 bp in length
20975 21074: gap of unknown length
21075 22409: contig of 1335 bp in length
22410 22509: gap of unknown length
22510 24337: contig of 1828 bp in length
24338 24438: gap of unknown length
24439 25826: contig of 1389 bp in length
25827 25926: gap of unknown length
25927 27769: contig of 1843 bp in length
27770 27869: gap of unknown length
27870 29441: contig of 1572 bp in length
29442 29541: gap of unknown length
29542 32494: contig of 2953 bp in length
32495 32594: gap of unknown length
32595 34633: contig of 2039 bp in length
34634 34733: gap of unknown length
34734 37159: contig of 2426 bp in length
37160 37259: gap of unknown length
37260 39230: contig of 1971 bp in length
39231 39330: gap of unknown length
39331 40642: contig of 1312 bp in length
40643 40742: gap of unknown length
40743 42839: contig of 1997 bp in length
42840 42839: gap of unknown length
42841 45245: contig of 2406 bp in length
45246 45345: gap of unknown length
45346 47083: contig of 1738 bp in length
47084 47183: gap of unknown length
47184 49945: contig of 2762 bp in length
49946 50045: gap of unknown length
50046 51544: contig of 1439 bp in length
51545 51644: gap of unknown length
51645 53891: contig of 2247 bp in length
53892 53991: gap of unknown length
53992 57321: contig of 3230 bp in length
57322 57321: gap of unknown length
57323 60016: contig of 2695 bp in length
60017 60116: gap of unknown length
60117 62563: contig of 2447 bp in length
```

```

* 62564 62663: gap of unknown length
* 62664 64926: contig of 2263 bp in length
* 64927 65026: gap of unknown length
* 65027 68373: contig of 3347 bp in length
* 68374 68473: gap of unknown length
* 68474 71700: contig of 3227 bp in length
* 71701 71800: gap of unknown length
* 71801 74355: contig of 2555 bp in length
* 74356 74456: gap of unknown length
* 74456 77267: contig of 2811 bp in length
* 77267 77367: gap of unknown length
* 77367 80106: contig of 2740 bp in length
* 80107 80206: gap of unknown length
* 80207 82763: contig of 2557 bp in length
* 82764 82863: gap of unknown length
* 82864 85878: contig of 3015 bp in length
* 85879 85978: gap of unknown length
* 85979 88748: contig of 2770 bp in length
* 88749 88848: gap of unknown length
* 88849 93166: contig of 4318 bp in length
* 93167 93266: gap of unknown length
* 93267 96634: contig of 3368 bp in length
* 96635 96734: gap of unknown length
* 96735 10049: contig of 3715 bp in length
* 10049 100549: gap of unknown length
* 100550 103083: contig of 2534 bp in length
* 103084 103183: gap of unknown length
* 103184 107029: contig of 3846 bp in length
* 107030 107129: gap of unknown length
* 107130 110324: contig of 3195 bp in length
* 110325 110424: gap of unknown length
* 110425 114667: contig of 4043 bp in length
* 114668 114567: gap of unknown length
* 114568 119345: contig of 4778 bp in length
* 119346 119445: gap of unknown length
* 119446 122439: contig of 2994 bp in length
* 122440 122539: gap of unknown length
* 122540 127598: contig of 5059 bp in length
* 127599 127698: gap of unknown length
* 127699 133792: contig of 6094 bp in length
* 133793 133892: gap of unknown length
* 133893 140225: contig of 6333 bp in length
* 140226 140325: gap of unknown length
* 140326 144336: contig of 4011 bp in length
* 144337 144436: gap of unknown length
* 144437 151171: contig of 6735 bp in length
* 151172 151271: gap of unknown length
* 151272 159310: contig of 8039 bp in length
* 159311 159410: gap of unknown length
* 159411 168312: contig of 8902 bp in length
* 168313 168412: gap of unknown length
* 168413 184375: contig of 15963 bp in length.

FEATURES
    source
        1.184375 Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="3"
            /map="3p"
            /clone="RP11-815J19"
            1.1389
            /note="assembly_name:Contig35"
            1490.3602
            /note="assembly_name:Contig37"
            3703.6147
            /note="assembly_name:Contig42"
            6248.7841
            /note="assembly_name:Contig45"
            7942.9569
            /note="assembly_name:Contig48"
            9670.11970
            /note="assembly_name:Contig50"
            12071.13209
            /note="assembly_name:Contig51"

```

```

misc_feature 13310..14906
              /note="assembly_name:Contig52"
misc_feature 15007..16915
              /note="assembly_name:Contig53"
misc_feature 17016..18655

```

```

Query Match 1.3%: Score 37; DB 2; Length 184375;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2462 attgagaacctgtctccattgtgtgtcttcctc 2498
      |||
Db 158390 ATTGAGAACCTGTCTCCATTGTGTGTCTTCCTC 158426

```

RESULT 14

AC027579

```

LOCUS AC027579 148270 bp DNA linear HTG 23-JUN-2000
DEFINITION Homo sapiens chromosome 17 clone RP11-177D1 map 17, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
AC027579
AC027579.2 GI:8671976
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 148270)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL
TITLE Homo sapiens chromosome 17, clone RP11-177D1
REFERENCE
2 (bases 1 to 148270)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakoque,K., Lamazares,R., Landers,T., Lenoczky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Meneus,L., Mihoval,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 23, 2000 this sequence version replaced gi:7342324.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9001
Center clone name: L77_D_1
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.96073i
Consensus quality: 141141 bases at least Q40

```

Consensus quality: 144792 bases at least Q30
Consensus quality: 146242 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 147070; sum-of-ctgigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-ctgigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 131: contig of 131 bp in length
132 231: gap of 100 bp
232 4256: contig of 4025 bp in length
4257 4356: gap of 100 bp
4357 6715: contig of 2359 bp in length
6716 6815: gap of 100 bp
6816 12854: contig of 6039 bp in length
12855 12954: gap of 100 bp
12955 18900: contig of 5946 bp in length
18901 19000: gap of 100 bp
19001 27560: contig of 8560 bp in length
27561 27660: gap of 100 bp
27661 41935: contig of 14275 bp in length
41936 42035: gap of 100 bp
42036 58175: contig of 16140 bp in length
58176 58275: gap of 100 bp
58276 70930: contig of 12655 bp in length
70931 71030: gap of 100 bp
71031 87632: contig of 16602 bp in length
87633 87732: gap of 100 bp
87733 104759: contig of 17027 bp in length
104760 104859: gap of 100 bp
104860 123863: contig of 19004 bp in length
123864 123963: gap of 100 bp
123964 148270: contig of 24307 bp in length.

FEATURES
source

1.148270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-177D1"
/clone_lib="RPCT-11 Human Male BAC"
1.131
/note="assembly_fragment"
clone_end:17
vector_side:right"
232.4256
/note="assembly_fragment"
4357.6715
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
6816.12854
/note="assembly_fragment"
12955.18900
/note="assembly_fragment"
19001.27560
/note="assembly_fragment"
27661.41935
/note="assembly_fragment"
42036.58175
/note="assembly_fragment"
58276.70930
/note="assembly_fragment"
71031.87632
/note="assembly_fragment"
87733.104759

misc_feature /note="assembly_fragment"
104860.123863
misc_feature /note="assembly_fragment"
123964.148270
BASE COUNT 40333 a 32099 c 32432 g 42206 t 1200 others
ORIGIN

Query Match 138: Score 36; DB 2; Length 148270;
Best Local Similarity 100.0%; Pred. No. 9; 1e-09;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2462 attgaacacgtcttcctccattgtgtgtcttct 2497
Db 76167 ATTGAGAACCGTCTTCCATTGCTGTCTTCTTCT 76202

RESULT 15
AC022727
LOCUS Homo sapiens chromosome 18 clone RP11-308J14 map 18, WORKING DRAFT
DEFINITION SEQUENCE, 11 unordered pieces.
AC022727
VERSION AC022727.4 GI:8072614
KEYWORDS HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 174405)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barina,N., Beckerly,R., Bede,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArillano,K., Dewar,K., Domino,M., Doyle,M., Fenebor,J.,
Ferrira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karalas,A., Klein,J.,
Landers,T., Lehocsky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McEneaney,K.,
McPheeters,R., Meldrum,J., Menes,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced g1:7658384.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5488
Center clone name: 308_J14
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169089 bases at least Q40
Consensus quality: 172164 bases at least Q30

Consensus quality: 173032 bases at least Q20
 Insert size: 177000; agarose-fp
 Insert size: 173405; sum-of-contigs
 Quality coverage: 4.8 in Q20 bases; agarose-fp
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 6450: contig of 6450 bp in length
 * 6451 6550: gap of 100 bp
 * 6551 12911: contig of 6361 bp in length
 * 12912 13011: gap of 100 bp
 * 13012 22057: contig of 9046 bp in length
 * 22058 22157: gap of 100 bp
 * 22158 32841: contig of 10684 bp in length
 * 32842 32941: gap of 100 bp
 * 32942 48367: contig of 15426 bp in length
 * 48368 48467: gap of 100 bp
 * 48468 63919: contig of 15452 bp in length
 * 63920 64019: gap of 100 bp
 * 64020 78259: contig of 14240 bp in length
 * 78260 78359: gap of 100 bp
 * 78360 98415: contig of 20056 bp in length
 * 98416 98515: gap of 100 bp
 * 98516 121876: contig of 23361 bp in length
 * 121877 121976: gap of 100 bp
 * 121977 144489: contig of 22513 bp in length
 * 144490 144589: gap of 100 bp
 * 144590 174405: contig of 29816 bp in length.

FEATURES
 Source
 1. 174405
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-308J14"
 /clone_lib="RPCT-11 Human Male BAC"
 1. 6450
 misc_feature
 /note="assembly_fragment"
 6551. 12911
 misc_feature
 /note="assembly_fragment"
 13012. 22057
 misc_feature
 /note="assembly_fragment"
 22158. 32841
 misc_feature
 /note="assembly_fragment"
 32942. 48367
 misc_feature
 /note="assembly_fragment"
 48468. 63919
 misc_feature
 /note="assembly_fragment"
 64020. 78259
 misc_feature
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right"
 78360. 98415
 misc_feature
 /note="assembly_fragment"
 clone_end:T7
 vector_side:left"
 98516. 121876
 misc_feature
 /note="assembly_fragment"
 121977. 144489
 misc_feature
 /note="assembly_fragment"
 144590. 174405
 misc_feature
 /note="assembly_fragment"
 BASE COUNT 46895 a 38223 c 37328 g 50956 t 1003 others
 ORIGIN

Query Match 1.3%; Score 36; DB 2; Length 174405;
 Best Local Similarity 100.0%; Pred. No. 9e-09;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2462 attggaacctgtctccattgtgtgtcttct 2497
 |||
 Db 127497 ATTGAGAACCTGTCTCCATTGTGTGTCTTCT 127532

RESULT 16

AP001531 179792 bp DNA linear HTG 08-AUG-2000
 LOCUS Homo sapiens chromosome 18 clone RP11-772F18 map 18p11.3, WORKING
 DEFINITION DRAFT SEQUENCE, 24 unordered pieces.
 AP001531
 VERSION AP001531.3 GI:9757457
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAT.
 SOURCE Homo sapiens DNA, clone:RP11-772F18.
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 179792)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE
 Homo sapiens 179,792 genomic DNA of 18p11.3
 JOURNAL
 Published Only in Database (2000) In press

AUTHORS
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (22-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
 E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-42-778-9923, Fax:81-42-778-9924
 On Aug 9, 2000 this sequence version replaced gi:8117372.

COMMENT

 1. 174405
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 Project Information
 Center clone name: RP11-772F18

 Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads Assembly
 Program: Phrap; version 0.990329
 Consensus quality: 171445 bases at least Q40
 Consensus quality: 173267 bases at least Q30
 Consensus quality: 175231 bases at least Q20
 Insert size: 177492; sum-of-contigs
 Quality coverage: 9.04x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of
 24 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved
 1 24375 contig of 24375 bp in length
 24476 45595 contig of 21120 bp in length
 45696 55640 contig of 9945 bp in length
 55741 70072 contig of 14332 bp in length
 70173 80776 contig of 10604 bp in length
 80877 89726 contig of 8849 bp in length
 89826 99984 contig of 10159 bp in length
 100085 108601 contig of 8517 bp in length
 108702 117836 contig of 9135 bp in length
 117937 125704 contig of 7768 bp in length
 125805 134648 contig of 9044 bp in length
 134949 142692 contig of 7744 bp in length

142793 151024 contig of 8232 bp in length
151125 157334 contig of 6210 bp in length
157435 162993 contig of 5559 bp in length
163094 167067 contig of 3974 bp in length
167168 169460 contig of 2293 bp in length
169561 171662 contig of 2102 bp in length
171763 173150 contig of 1388 bp in length
173251 175114 contig of 1864 bp in length
175215 176405 contig of 1191 bp in length
175506 177563 contig of 1058 bp in length
177654 178673 contig of 1010 bp in length
178774 179792 contig of 1019 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 24375: contig of 24375 bp in length
24376 24475: gap of 100 bp
24476 45595: contig of 21120 bp in length
45596 45695: gap of 100 bp
45696 55640: contig of 9945 bp in length
55641 55740: gap of 100 bp
55741 70072: contig of 14332 bp in length
70073 70172: gap of 100 bp
70173 80776: contig of 10604 bp in length
80777 80876: gap of 100 bp
80877 89725: contig of 8849 bp in length
89726 89825: gap of 100 bp
89826 99984: contig of 10159 bp in length
99985 100084: gap of 100 bp
100085 108601: contig of 8517 bp in length
108602 108701: gap of 100 bp
108702 117836: contig of 9135 bp in length
117837 117936: gap of 100 bp
117937 125704: contig of 7768 bp in length
125705 125804: gap of 100 bp
125805 134848: contig of 9044 bp in length
134849 134948: gap of 100 bp
134949 142692: contig of 7744 bp in length
142693 142792: gap of 100 bp
142793 151024: contig of 8232 bp in length
151025 151124: gap of 100 bp
151125 157334: contig of 6210 bp in length
157335 157434: gap of 100 bp
157435 162993: contig of 5559 bp in length
162994 163093: gap of 100 bp
163094 167067: contig of 3974 bp in length
167068 167167: gap of 100 bp
167168 169460: contig of 2293 bp in length
169461 169560: gap of 100 bp
169561 171662: contig of 2102 bp in length
171663 171762: gap of 100 bp
171763 173150: contig of 1388 bp in length
173151 173250: gap of 100 bp
173251 175114: contig of 1864 bp in length
175115 175214: gap of 100 bp
175215 176405: contig of 1191 bp in length
176406 176505: gap of 100 bp
176506 177563: contig of 1058 bp in length
177564 178673: contig of 1010 bp in length
178674 178773: gap of 100 bp
178774 179792: contig of 1019 bp in length.

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p11.3"

/clone="RP11-772F18"
1. .24375
/note="assembly-fragment"
24476..45595
/note="assembly-fragment"
45696..55640
/note="assembly-fragment"
55741..70072
/note="assembly-fragment"
70173..80776
/note="assembly-fragment"
80877..89725
/note="assembly-fragment"
89826..99984
/note="assembly-fragment"
100085..108601
/note="assembly-fragment"
108702..117836
/note="assembly-fragment"
117937..125704
/note="assembly-fragment"
125805..134848
/note="assembly-fragment"
134949..142692
/note="assembly-fragment"
142793..151024
/note="assembly-fragment"
151125..157334
/note="assembly-fragment clone_end:SP6 vector_side:right"
157435..162993
/note="assembly-fragment"
163094..167067
/note="assembly-fragment"
167168..169460
/note="assembly-fragment clone_end:T7 vector_side:left"
169561..171662
/note="assembly-fragment"
171763..173150
/note="assembly-fragment"
173251..175114
/note="assembly-fragment"
175215..176405
/note="assembly-fragment"
176506..177563
/note="assembly-fragment"
177664..178673
/note="assembly-fragment"
178774..179792
/note="assembly-fragment"
BASE COUNT 49430 a 37736 c 39365 g 50954 t 2307 others
ORIGIN

Query Match 1.3%: Score 36; DB 2; Length 179792;
Best Local Similarity 100.0%; Pred. No. 9e-09;

Matches .36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 atgagacgtctccattgctgtcttcct 2497
|||||
Db 29905 ATTGAGACCTGCTTCCATTGCTGCTTCC 29940

RESULT 17

AC025211

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

150794 bp DNA linear PRI 08-SEP-2001
Homo sapiens chromosome , clone RP11-220C2, complete sequence.
HTG.
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.


```

repeat_region      /rpt_family="L2"
                    complement(17671..17984)
                    /rpt_family="AluJb"
repeat_region      /rpt_family="L1B1"
                    complement(18115..18681)
repeat_region      /rpt_family="L1B1"
                    complement(18783..19023)
repeat_region      /rpt_family="L1B1"
                    complement(19024..19321)
repeat_region      /rpt_family="AluSg"
                    complement(19322..20092)
repeat_region      /rpt_family="L1B1"
                    complement(20098..20194)
repeat_region      /rpt_family="Cr-rich"
                    complement(20195..20356)
repeat_region      /rpt_family="AluJb"
                    complement(20357..20395)
repeat_region      /rpt_family="TGCn"
                    complement(20396..20690)
repeat_region      /rpt_family="AluSx"
                    complement(20691..20830)
repeat_region      /rpt_family="AluJb"
                    complement(20831..22342)
repeat_region      /rpt_family="L1B1"
                    complement(21587..21602)
unsure             /note="<30 qual SNGU region"
                    complement(21644..21658)
unsure             /note="<30 qual SNGU region"
                    complement(22343..22535)
repeat_region      /rpt_family="AluJb"
                    complement(22629..22718)
repeat_region      /rpt_family="AluJ/FLAM"
                    complement(23036..23116)
repeat_region      /rpt_family="L1B2"
                    complement(23117..23399)
unsure             /rpt_family="AluJo"
                    complement(23294..23299)
repeat_region      /note="<30 qual SNGU region"
                    complement(23400..23422)
repeat_region      /rpt_family="L1B2"
                    complement(23472..23561)
repeat_region      /rpt_family="MIR"
                    complement(24063..24374)
repeat_region      /rpt_family="AluSx"
                    complement(24989..25192)
repeat_region      /rpt_family="L3"

Query Match      1.2%: Score 34: DB 9: Length 150794:
Best Local Similarity 100.0%: Pred No. 1.3e-07:
Matches 34: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 2464 tgaagacctgtctccattgtgtgtcttctt 2497
      |||||
Db 109211 TGAGAACCTGCTCCATTGTGTCCTTCC 109244

RESULT 18
AC087493 166650 bp DNA linear HTG 24-MAY-2001
LOCUS AC087493/c
DEFINITION Homo sapiens chromosome 17 clone RP11-253N19 map 17, WORKING DRAFT
ACCESSION AC087493
VERSION AC087493.3 GI:14192363
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 166650)
REFERENCE 1 (bases 1 to 166650)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-253N19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 166650)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

```

TITLE JOURNAL COMMENT

```

Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Camarate,J., Campopiano,A., Chepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hages,B., Heatford,A., Horton,L., Hulme,W., Illey,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Jehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPhyeters,R., Meldrim,J., Meneus,L., Minova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thmann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Treviers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zalnoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2001 this sequence version replaced gi:13123255.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L11973
Center clone name: 253_N_19

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165769 bases at least Q40
Consensus quality: 166122 bases at least Q30
Consensus quality: 166247 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 166350; sum-of-ctnigs
Quality coverage: 11.0 in Q20 bases; agarose-fp
Quality coverage: 10.8 in Q20 bases; sum-of-ctnigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11793: contig of 11793 bp in length
* 11794 11893: gap of 100 bp
* 11894 33667: contig of 21774 bp in length
* 33668 33767: gap of 100 bp
* 33768 97497: contig of 63730 bp in length
* 97498 97597: gap of 100 bp
* 97598 166650: contig of 69053 bp in length.
*
* Location/Qualifiers
* 1. 166650
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="17"
* /map="17"
* /clone="RP11-253N19"
* /clone_1lb="RP11-11 Human Male BAC"
* 1. 11793
* /note="assembly-fragment
* clone_end:SP6

```

misc_feature vector_side:left"
11894..33667
/note="assembly_fragment"
misc_feature 33768..97497
/note="assembly_fragment"
misc_feature 97598..166650
/note="assembly_fragment"
clone_end:17
vector_side:right"

BASE COUNT 51731 a 33076 c 32370 g 49168 t 305 others
ORIGIN

Query Match 1.2%; Score 34; DB 2; Length 166650;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2464 ttgagaccgtctccatctgtgtgtcttctc 2497
Db 39715 ttgagAACCTGCTCCATTTGCTGCTTCC 39682

RESULT 19
AC008106 218922 bp DNA linear HTG 04-SEP-2000
LOCUS Homo sapiens chromosome 17 clone CTD-3008D3 map 17, WORKING DRAFT
DEFINITION
AC008106
AC008106
AC008106.3 GI:9966323
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
human.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 218922)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, M., Forrest, C., Funes, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Headford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczeky, J., Liu, C., Locke, K., Macdonald, P., Margulis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Melidim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Testaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
Submitted (22-JUL-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 4, 2000 this sequence version replaced gi:7329486.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 3008_D_3
Center clone name: 1900
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-primer; amersham; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 197177 bases at least Q40
Consensus quality: 206822 bases at least Q30
Consensus quality: 211643 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 215622; sum-of-contigs
Quality coverage: 4.1 in Q20 bases; agarose-fp
Quality coverage: 3.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1090: contig of 1090 bp in length
1091 1190: gap of 100 bp
1191 2690: contig of 1500 bp in length
2691 2790: gap of 100 bp
2791 4012: contig of 1222 bp in length
4013 4112: gap of 100 bp
4113 5123: contig of 1011 bp in length
5124 5223: gap of 100 bp
5224 6312: contig of 1089 bp in length
6313 6412: gap of 100 bp
6413 8474: contig of 2062 bp in length
8475 8574: gap of 100 bp
8575 10328: contig of 1753 bp in length
10329 10428: gap of 100 bp
10429 12116: contig of 1688 bp in length
12117 12216: gap of 100 bp
12217 14183: contig of 1967 bp in length
14184 14283: gap of 100 bp
14284 17330: contig of 3047 bp in length
17331 17430: gap of 100 bp
17431 21134: contig of 3704 bp in length
21135 21234: gap of 100 bp
21235 24428: contig of 3194 bp in length
24429 24528: gap of 100 bp
24529 27705: contig of 3177 bp in length
27706 27805: gap of 100 bp
27806 30106: contig of 2301 bp in length
30107 30206: gap of 100 bp
30207 35366: contig of 5160 bp in length
35367 35466: gap of 100 bp
35467 39634: contig of 4168 bp in length
39635 39734: gap of 100 bp
39735 43767: contig of 4033 bp in length
43768 43867: gap of 100 bp
43868 50104: contig of 6237 bp in length
50105 50204: gap of 100 bp
50205 54896: contig of 4692 bp in length
54897 54996: gap of 100 bp
54997 61677: contig of 6681 bp in length
61678 61777: gap of 100 bp
61778 68042: contig of 6265 bp in length
68043 68142: gap of 100 bp
68143 76003: contig of 7861 bp in length
76004 76103: gap of 100 bp
76104 83935: contig of 7832 bp in length
83936 84035: gap of 100 bp
84036 92731: contig of 8696 bp in length
92732 92831: gap of 100 bp
92832 102205: contig of 9374 bp in length
102206 102305: gap of 100 bp
102306 111555: contig of 9250 bp in length
111556 111655: gap of 100 bp
111656 120772: contig of 9117 bp in length
120773 120872: gap of 100 bp
120873 129869: contig of 8997 bp in length
129870 129969: gap of 100 bp
129970 143160: contig of 13191 bp in length

```
* 143161 143260: gap of 100 bp
* 143261 154541: contig of 11281 bp in length
* 154542 154641: gap of 100 bp
* 154642 170485: contig of 15844 bp in length
* 170486 170585: gap of 100 bp
* 170586 183973: contig of 13388 bp in length
* 183974 184073: gap of 100 bp
* 184074 200761: contig of 16688 bp in length
* 200762 200861: gap of 100 bp
* 200862 218922: contig of 18061 bp in length.

FEATURES
    source
        1..218922
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="17"
            /map="17"
            /clone="CTD-3008D3"
            /clone_1lb="CTD Human BAC"
            1..1090
                /note="assembly_fragment"
            1191..2690
                /note="assembly_fragment"
            2791..4012
                /note="assembly_fragment"
            4113..5123
                /note="assembly_fragment"
            5224..6312
                /note="assembly_fragment"
            6413..8474
                /note="assembly_fragment"
            8575..10328
                /note="assembly_fragment"
            10429..12116
                /note="assembly_fragment"
            12217..14183
                /note="assembly_fragment"
            14284..17330
                /note="assembly_fragment"
            17431..21134
                /note="assembly_fragment"
            21235..24428
                /note="assembly_fragment"
            24529..27705
                /note="assembly_fragment"
            27806..30106
                /note="assembly_fragment"
            30207..35366
                /note="assembly_fragment"
            35467..39634
                /note="assembly_fragment"
            39735..43767
                /note="assembly_fragment"
            43868..50104
                /note="assembly_fragment"
            50205..54896
                /note="assembly_fragment"
            54997..61677
                /note="assembly_fragment"
            61778..68042
                /note="assembly_fragment"
            68143..76003
                /note="assembly_fragment"
            76104..83935
                /note="assembly_fragment"
            84036..92731
                /note="assembly_fragment"
            92832..102205
                /note="assembly_fragment"
            102306..111555
                /note="assembly_fragment"
            111656..120772
                /note="assembly_fragment"
            120873..129869
                /note="assembly_fragment"
```

```
misc_feature      /note="assembly_fragment"
                  129970..143160
misc_feature      /note="assembly_fragment"
                  143261..154541
misc_feature      /note="assembly_fragment"

Query Match      1.2%: Score 34; DB 2; Length 218922;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2464 tgaagacctgtctccattgtgtgtcttct 2497
Db 99714 TGAGAACCTGTCCTCCATTGCTGTCCTCC 99747

RESULT 20
AC027607/c
LOCUS AC027607 162453 bp DNA linear HTG 31-AUG-2001
DEFINITION Homo sapiens chromosome 4 clone RP11-36B15, WORKING DRAFT SEQUENCE,
3 unordered pieces.
ACCESSION AC027607
VERSION AC027607.6 GI:15321552
KEYWORDS HTG; HTGS_PHASE1; HTGS_DNAF1; HTGS_ACTIVIERIN.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 162453)
JOURNAL The sequence of Homo sapiens clone
AUTHORS Unpublished
TITLE 2 (bases 1 to 162453)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (30-MAR-2000) Genome Sequencing Center, Washington
AUTHORS University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
COMMENT On Aug 28, 2001 this sequence version replaced gi:15145259.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0036B15

----- Summary Statistics -----
Sequencing vector: M13; 79%
Sequencing vector: plasmid; 21%
Chemistry: Dye-primer ET; 79% of reads
Chemistry: Dye-terminator Big Dye; 21% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162283 bases at least Q40
Consensus quality: 162660 bases at least Q30
Consensus quality: 162828 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 163582; sum-of-ctnigs
Quality coverage: 9.33 in Q20 bases; agarose-fp
Quality coverage: 8.80 in Q20 bases; sum-of-ctnigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1192: contig of 1192 bp in length
* 1193 1292: gap of unknown length
* 1293 72574: contig of 71282 bp in length
* 72575 72674: gap of unknown length
* 72675 162453: contig of 89779 bp in length.
```

FEATURES
source
Location/Qualifiers
1.162453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-36B15"
1.1192
/note="assembly_name:Contig6"
misc_feature
1293..72574
/note="assembly_name:Contig10"
72675..162453
/note="assembly_name:Contig11
clone_end:T7
vector_side:left"
BASE COUNT 46434 a 33274 c 34864 g 47681 t 200 others
ORIGIN
Query Match 1.2%; Score 32; DB 2; Length 162453;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2362 agtgcaggagcagtcacatgccagcagcaga 2393
|||||
Db 96257 AGTGCAGGAGGACGACGATGCAATGCCAGCAGCAGA 96226
|||||
RESULT 21
AC036185/c 171279 bp DNA linear HTG 21-JUN-2000
LOCUS Homo sapiens chromosome 4 clone RP11-679B6 map 4, WORKING DRAFT
DEFINITION
SEQUENCE, 15 unordered pieces.
AC036185
AC036185.2 GI:8576273
HTG: HTGS_PHA5B1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 171279)
Britten, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 4, clone RP11-679B6
Unpublished
2 (bases 1 to 171279)
Britten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,
Boguslavsky, L., Bouckgealter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choe, P., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S.,
Dodgson, S., Dominko, M., Doyle, M., Ferreira, P., FitzHugh, M., Gage, D.,
Galagan, J., Gardina, S., Glend, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heath, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A.,
Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPheters, R.,
Meldrum, J., Menees, L., Mihov, T., Miranda, C., Mlepta, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tejeda, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9247
Center clone name: 679_B-6
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 163254 bases at least Q40
Consensus quality: 166954 bases at least Q30
Consensus quality: 168680 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 169879; sum-of-ctigs
Quality coverage: 7.1 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum-of-ctigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1385: contig of 1385 bp in length
1386 1485: gap of 100 bp
1486 2835: contig of 1350 bp in length
2836 2935: gap of 100 bp
2936 4108: contig of 1173 bp in length
4109 4208: gap of 100 bp
4209 7507: contig of 3299 bp in length
7508 7607: gap of 100 bp
7608 13863: contig of 6262 bp in length
13870 13969: gap of 100 bp
13970 20810: contig of 6841 bp in length
20811 20910: gap of 100 bp
20911 28893: contig of 7883 bp in length
28894 28993: gap of 100 bp
28994 33086: contig of 10093 bp in length
33087 39186: gap of 100 bp
39187 50077: contig of 10891 bp in length
50078 50177: gap of 100 bp
50178 61526: contig of 11349 bp in length
61527 61625: gap of 100 bp
61627 75523: contig of 13903 bp in length
75530 75629: gap of 100 bp
75630 91544: contig of 15915 bp in length
91545 91644: gap of 100 bp
91645 108750: contig of 17106 bp in length
108751 108850: gap of 100 bp
108851 136877: contig of 28027 bp in length
136878 136977: gap of 100 bp
136978 171279: contig of 34302 bp in length.
Location/Qualifiers
1.171279
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-679B6"
/clone_id="RPCT-11 Human Male BAC"
1.1385
/note="assembly_fragment"
1486..2835
/note="assembly_fragment"
2936..4108
/note="assembly_fragment"
4209..7507
/note="assembly_fragment"
7608..13869

```

misc_feature /note="assembly_fragment"
13970..20810
/note="assembly_fragment"
misc_feature /note="assembly_fragment"
20911..28893
/note="assembly_fragment"
misc_feature 28994..39086
/note="assembly_fragment"
misc_feature 39187..50077
/note="assembly_fragment"
misc_feature 50178..61526
/note="assembly_fragment"
misc_feature 61627..75529
/note="assembly_fragment"
clone_end:T7
vector_side:right"
75630..91544
/note="assembly_fragment"
91645..108750
/note="assembly_fragment"
108851..136877
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
136978..171279
/note="assembly_fragment"
BASE COUNT 48027 a 36049 c 36303 g 49499 t 1401 others
ORIGIN

```

```

Query Match 1.2%; Score 32; DB 2; Length 171279;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2362 agtgcgagagagcagcgaatgccgagcaga 2393
|||||
Db 143009 AGTGCAGAGCAGTCATATGCCAGCAGCA 142978

```

```

RESULT 22
AC026269 171601 bp DNA 1linear HTG 06-DEC-2001
LOCUS Homo sapiens chromosome 11 clone RP11-745E9 map 11, *** SEQUENCING
IN PROGRESS ***, 5 unordered pieces.
ACCESSION AC026269
VERSION AC026269.6 GI:17386474
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 171601)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-745E9
Unpublished
2 (bases 1 to 171601)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campanaro,A., Castle,A., Choepel,Y., Colaneri,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehotzky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGarr,A., McKernan,K., McNeelers,R.,
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Miengo,A., Morrow,J.,
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

```

```

TITLE Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody.M.
JOURNAL Direct Submission
COMMENT Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 6, 2001 this sequence version replaced g1:13959308.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 18600
Center clone name: 745_E_9

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 28797: contig of 28797 bp in length
* 28798 28897: gap of 100 bp
* 28898 52568: contig of 23671 bp in length
* 52569 52668: gap of 100 bp
* 52669 104066: contig of 51398 bp in length
* 104067 104166: gap of 100 bp
* 104167 140470: contig of 36304 bp in length
* 140471 140570: gap of 100 bp
* 140571 171601: contig of 31031 bp in length.
Location/Qualifiers
1..171601
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-745E9"
/clone_lib="RPC1-11 Human Male BAC"

```

```

BASE COUNT 43801 a 37206 c 38797 g 51317 t 480 others
ORIGIN

```

```

Query Match 1.2%; Score 32; DB 2; Length 171601;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2703 cctgcacccctccactgagagctgtct 2734
|||||
Db 89086 CCTGCATCCCTCTCCACTGAGAGCTGTCT 89055

```

```

RESULT 23
AC016166 173133 bp DNA 1linear HTG 20-SEP-2000
LOCUS Homo sapiens clone RP11-1D6, WORKING DRAFT SEQUENCE, 31 unordered
pieces.
ACCESSION AC016166
VERSION AC016166.3 GI:10196377
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 173133)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1D6
Unpublished

```

REFERENCE
AUTHORS

2 (bases 1 to 173133)
Bairden, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donellan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galand, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lechoczky, J., Lien, C., Locke, K., Macdonald, P., Margolis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, T., Meltrin, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (23-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 20, 2000 this sequence version replaced gi:108714.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project information
Center project name: L1434
Center clone name: 1_D_6

Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156896 bases at least Q40
Consensus quality: 165527 bases at least Q40
Consensus quality: 168595 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 170133; sum-of-contigs
Quality coverage: 3.5 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 5114: contig of 5114 bp in length
* 5115 5214: gap of 100 bp
* 5215 6404: contig of 1190 bp in length
* 6405 6504: gap of 100 bp
* 6505 7651: contig of 1147 bp in length
* 7652 7751: gap of 100 bp
* 7752 9355: contig of 1604 bp in length
* 9356 9455: gap of 100 bp
* 9456 10655: contig of 1200 bp in length
* 10656 10755: gap of 100 bp
* 10756 11937: contig of 1182 bp in length
* 11938 12037: gap of 100 bp
* 12038 13055: contig of 1018 bp in length
* 13056 13155: gap of 100 bp
* 13156 14204: contig of 1049 bp in length
* 14205 14304: gap of 100 bp
* 14305 15800: contig of 1496 bp in length
* 15801 15900: gap of 100 bp
* 15901 17319: contig of 1419 bp in length
* 17320 17419: gap of 100 bp
* 17420 19622: contig of 2203 bp in length
* 19623 19722: gap of 100 bp
* 19723 21692: contig of 1970 bp in length

FEATURES
source

misc_feature
21693 21792: gap of 100 bp
* 21793 23722: contig of 1930 bp in length
* 23723 23822: gap of 100 bp
* 23823 27350: contig of 3528 bp in length
* 27351 27450: gap of 100 bp
* 27451 30128: contig of 2678 bp in length
* 30129 30228: gap of 100 bp
* 30229 33435: contig of 3207 bp in length
* 33436 33535: gap of 100 bp
* 33536 36762: contig of 3227 bp in length
* 36763 36862: gap of 100 bp
* 36863 41461: contig of 4599 bp in length
* 41462 41561: gap of 100 bp
* 41562 45986: contig of 4425 bp in length
* 45987 46086: gap of 100 bp
* 46087 51166: contig of 5080 bp in length
* 51167 51266: gap of 100 bp
* 51267 55890: contig of 4624 bp in length
* 55891 55990: gap of 100 bp
* 55991 62876: contig of 6886 bp in length
* 62877 62976: gap of 100 bp
* 62977 70430: contig of 7454 bp in length
* 70431 70530: gap of 100 bp
* 70531 87882: contig of 17352 bp in length
* 87883 87982: gap of 100 bp
* 87983 97567: contig of 9585 bp in length
* 97568 97667: gap of 100 bp
* 97668 106038: contig of 8371 bp in length
* 106039 106138: gap of 100 bp
* 106139 118960: contig of 12822 bp in length
* 118961 119060: gap of 100 bp
* 119061 133793: contig of 14733 bp in length
* 133794 133893: gap of 100 bp
* 133894 148302: contig of 14809 bp in length
* 148303 148402: gap of 100 bp
* 148403 165690: contig of 17288 bp in length
* 165691 16590: gap of 100 bp
* 16591 173133: contig of 7343 bp in length.
* 165791 Location/Qualifiers

1. 173133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-106"
/clone_lib="RP11 Human Male BAC"
1. 5114
/note="assembly_fragment"
clone_end:SP6
vector_side:left
5215..6404
/note="assembly_fragment"
6505..7651
/note="assembly_fragment"
7752..9355
/note="assembly_fragment"
9456..10655
/note="assembly_fragment"
10756..11937
/note="assembly_fragment"
12038..13055
/note="assembly_fragment"
13156..14204
/note="assembly_fragment"
14305..15800
/note="assembly_fragment"
15901..17319
/note="assembly_fragment"
17420..19622
/note="assembly_fragment"
19723..21692
/note="assembly_fragment"
21793..23722
/note="assembly_fragment"
23823..27350

```

misc_feature      /note="assembly_fragment"
27451..30128
/note="assembly_fragment"
misc_feature      30229..33435
/note="assembly_fragment"
misc_feature      33536..36762
/note="assembly_fragment"
misc_feature      36863..41461
/note="assembly_fragment"
misc_feature      41562..45986
/note="assembly_fragment"
misc_feature      46087..51166
/note="assembly_fragment"
misc_feature      51267..55890
/note="assembly_fragment"
misc_feature      55991..62876
/note="assembly_fragment"
misc_feature      62977..70430
/note="assembly_fragment"
misc_feature      70531..87882
/note="assembly_fragment"
misc_feature      87983..97567
/note="assembly_fragment"
misc_feature      97668..106038
/note="assembly_fragment"
misc_feature      106139..118960
/note="assembly_fragment"
misc_feature      119061..133793
/note="assembly_fragment"
misc_feature      133894..148302
/note="assembly_fragment"
misc_feature      148403..165690
/note="assembly_fragment"
misc_feature      165791..173133
/note="assembly_fragment"
misc_feature      clone_end:T7
vector_side:right"
BASE COUNT      51650 a 33587 c 33772 g 51116 t 3008 others
ORIGIN

```

```

Query Match      1.2%: Score 32; DB 2; Length 173133;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2457 actgaattggaacctgtctccattggtg 2488
|||||
Db 15324 ACTGAATTGAGACCTCTCTCCATTGGTG 15293

```

```

RESULT 24
AC098581          174701 bp   DNA      linear   HTG 25-OCT-2001
LOCUS            Homo sapiens chromosome 2 clone RP11-1D6, WORKING DRAFT SEQUENCE, 9
DEFINITION      unordered pieces.
ACCESSION      AC098581
VERSION        AC098581.1 GI:16418216
KEYWORDS        HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 174701)
AUTHORS        Waterston,R.H.
TITLE          The sequence of Homo sapiens clone
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 174701)
AUTHORS        Waterston,R.H.
TITLE          Direct Submission
JOURNAL        Submitted (25-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H_NH0001D06
----- Summary Statistics -----
Sequencing vector: MJ3: 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171641 bases at least Q40
Consensus quality: 172606 bases at least Q30
Consensus quality: 173237 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1406: contig of 1406 bp in length
* 1407 1506: gap of unknown length
* 1507 6809: contig of 5303 bp in length
* 6810 6909: gap of unknown length
* 6910 20466: contig of 13557 bp in length
* 20467 20566: gap of unknown length
* 20567 35544: contig of 14978 bp in length
* 35545 35645: gap of unknown length
* 35645 52465: contig of 16821 bp in length
* 52466 52565: gap of unknown length
* 52566 73259: contig of 20654 bp in length
* 73260 73359: gap of unknown length
* 73360 105109: contig of 31750 bp in length
* 105110 105209: gap of unknown length
* 105210 140153: contig of 34944 bp in length
* 140154 140253: gap of unknown length
* 140254 174701: contig of 34448 bp in length.
Location/Qualifiers
1. 174701
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-1D6"
1. 1406
/note="assembly_name:Contig5"
1507..6809
/note="assembly_name:Contig6"
6910..20466
/note="assembly_name:Contig7"
20567..35544
/note="assembly_name:Contig8"
35645..52465
/note="assembly_name:Contig9
clone_end:T7
vector_side:right"
52566..73259
/note="assembly_name:Contig10"
73360..105109
/note="assembly_name:Contig11"
105210..140153
/note="assembly_name:Contig12"
140254..174701
/note="assembly_name:Contig13"
BASE COUNT      52648 a 34002 c 34946 g 52304 t 801 others
ORIGIN

```



```

misc_feature      75817..84841
                  /note="assembly_name:Contig34"
misc_feature      84942..95395
                  /note="assembly_name:Contig35"
misc_feature      95496..110306
                  /note="assembly_name:Contig36"
misc_feature      110407..123845
                  /note="assembly_name:Contig37"
misc_feature      123946..142399
                  /note="assembly_name:Contig38"
misc_feature      142500..157956
                  /note="assembly_name:Contig39"
misc_feature      158057..173833
                  /note="assembly_name:Contig40"
BASE COUNT      50115 a 38599 c 38438 g 46468 t 2213 others
ORIGIN

```

```

Query Match      1.28; Score 32; DB 2; Length 175833;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2703 cccgcacccctccactgagagcgtgtct 2734
      ||||||||||||||||||||||||||||
Db 66179 CCCTGCAHCCCTCTCCACTGAGAGCTGTCT 66210

```

```

RESULT 26
AC007380/c      207362 bp      DNA      linear      PRI 29-APR-2000
LOCUS
DEFINITION      Homo sapiens BAC clone RP11-150013 from 2, complete sequence.
ACCESSION      AC007380
VERSION      AC007380.3 GI:5732167
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 207362)
Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
REFERENCE
AUTHORS      Du, F., Maupin, R. and Sowa, D.
2 (bases 1 to 207362)
The sequence of Homo sapiens BAC clone RP11-150013
JOURNAL      Unpublished
REFERENCE
AUTHORS      Waterston, R.H.
3 (bases 1 to 207362)
Direct Submission
JOURNAL      Submitted (25-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 207362)
Waterston, R.H.
JOURNAL      Submitted (13-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 207362)
Waterston, R.
JOURNAL      Direct Submission
Submitted (29-APR-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 13, 1999 this sequence version replaced gi:5001523.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
-----
Summary Statistics
-----
Center project name: H_NH0150013

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatem, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-355P16. Actual start of this clone is at base position 1 of RP11-150013; actual end is at base position 207362 of RP11-150013.

FEATURES

```

source
1..207362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-150013"
/clone_1lb="RPCI-11"
13..437
/rpt_family="L1"
repeat_region
438..738
/rpt_family="Alu"
739..1003
/rpt_family="L1"
1160..2723
/rpt_family="L1"
2911..3114
/rpt_family="MERL_type"
3727..3773
/rpt_family="(TTTA)n"
3775..3975
/rpt_family="L2"
5146..5215
/rpt_family="(TA)n"
5848..6147
/rpt_family="Alu"
8227..8472
/rpt_family="Alu"
8876..9591
/rpt_family="L1"
9922..10240
/rpt_family="Other"
11220..11451
/rpt_family="Retroviral"
11784..11949
/rpt_family="MIR"
12040..12079
/rpt_family="L2"
12095..12288
repeat_region

```

```

repeat_region      /rpt_family="Retroviral"
                    12788..13196
repeat_region      /rpt_family="MaLR"
                    13360..13408
repeat_region      /rpt_family="L2"
                    13432..13581
repeat_region      /rpt_family="L2"
                    13624..13921
repeat_region      /rpt_family="Alu"
                    13986..15733
repeat_region      /rpt_family="L1"
                    15734..15777
repeat_region      /rpt_family="Retroviral"
                    15948..17035
repeat_region      /rpt_family="L1"
                    17134..18265
repeat_region      /rpt_family="L1"
                    18932..19021
repeat_region      /rpt_family="MER81"
                    19038..19270
repeat_region      /rpt_family="MIR"
                    19430..20792
repeat_region      /rpt_family="L2"
                    20808..20859
repeat_region      /rpt_family="MER1_type"
                    20889..22563
repeat_region      /rpt_family="L1"
                    25258..25560
repeat_region      /rpt_family="Alu"
                    25642..25793
repeat_region      /rpt_family="MER1_type"
                    25825..25855
repeat_region      /rpt_family="AT_rich"
                    26170..26882
repeat_region      /rpt_family="L1"
                    26883..27043
repeat_region      /rpt_family="Retroviral"
                    27091..27399
repeat_region      /rpt_family="Retroviral"
                    27401..28719
repeat_region      /rpt_family="L1"
                    28720..29042
repeat_region      /rpt_family="Alu"
                    29043..30738
repeat_region      /rpt_family="L1"
                    30777..31443
repeat_region      /rpt_family="L1"
                    31822..32278
repeat_region      /rpt_family="L1"
                    32588..33491
repeat_region      /rpt_family="L2"
                    33500..33681
repeat_region      /rpt_family="MIR"
                    33891..33949
repeat_region      /rpt_family="GA-rich"
                    35130..35210
repeat_region      /rpt_family="MIR"
                    35296..35652
repeat_region      /rpt_family="MaLR"
                    35682..35841
repeat_region      /rpt_family="MIR"
                    35852..36079
repeat_region      /rpt_family="L2"
                    36413..36484
repeat_region      /rpt_family="Mariner"
                    37368..37399
repeat_region      /rpt_family="AT_rich"
                    37607..37780
repeat_region      /rpt_family="Other"
                    37940..38176
repeat_region      /rpt_family="L1"
                    38193..38312
repeat_region      /rpt_family="Other"

```

```

repeat_region      38520..38872
                    /rpt_family="Other"
repeat_region      38873..39166
                    /rpt_family="Alu"
repeat_region      39167..39243
                    /rpt_family="Other"
repeat_region      39244..39290
                    /rpt_family="(TRCA)n"
repeat_region      39291..39494
                    /rpt_family="Other"

Query Match      1.28; Score 32; DB 9; Length 207362;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2457 actgaattgagaacctgtctccattggcg 2488
Db 183210 ACTGAATTGAGAACCTGTCTCCATTGGGTG 183179
|||||

```

```

RESULT 27
LOCUS      AL157944      154746 bp      DNA      linear      PRI 19-JAN-2001
DEFINITION Human DNA sequence from clone RP5-1126B2 on chromosome 1p31.1-31.2,
complete sequence.
ACCESSION      AL157944
VERSION      AL157944.24 GI:12329215
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 154746)
Lovel1.J.
Direct Submission
Submitted (19-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequest@sanger.ac.uk
On Jan 21, 2001 this sequence version replaced gi:12227321.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM: EMBL; SW: SWISSPROT; TR: TrEMBL; WP: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-1126B2 is from the library RPCR-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/Bacpac/home.htm
VECTOR: pCYPAC2

```

This sequence is the entire insert of clone RP5-1126B2 The true left end of clone RP4-600K12 is at 95393 in this sequence. The true left end of clone RP11-170N11 is at 125881 in this sequence. The true right end of clone RP11-14M11 is at 52363 in this sequence.

```

FEATURES
    source
        1..154746
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="p31.1-31.2"

```

```

misc_feature /clone="RP5-1126B2"
              /clone_id="RPCL-5"
              6468..6551
              /note="Sequence from overlapping clone ba14M11(AL356974)
              Assembly confirmed by restriction digest"
misc_feature 65174..65189
              /note="Single clone region. Assembly confirmed by
              restriction digest data"
misc_feature 65516..65548
              /note="Single clone region. Assembly confirmed by
              restriction digest data"
misc_feature 108104..108297
              /note="Sequence from overlapping clone dj600K12(AL357142)
              Assembly confirmed by restriction digest"
misc_feature 126203..126627
              /note="Sequence from overlapping clones dj600K12(AL357142)
              and ba170N11(AL354654) Assembly confirmed by restriction
              digest"
misc_feature 126429..126627
              /note="Sequence from overlapping clones dj600K12(AL357142)
              and ba170N11(AL354654) Assembly confirmed by restriction
              digest"
BASE COUNT 49641 a 27937 c 28135 g 49033 t
ORIGIN

```

```

Query Match 1.1%; Score 31; DB 9; Length 154746;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2462 atgagaacctgtctccattgtgtgt 2492
Db 53581 ATTGAGACCTGTCTCCATTGTGTGT 53611

```

```

RESULT 28
AC019042/c 168736 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 8 clone RP11-11A18, WORKING DRAFT SEQUENCE,
DEFINITION 5 unordered pieces.
ACCESSION AC019042
VERSION AC019042.3 GI:7243881
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT;
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 168736)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 168736)
TITLE Waterston,R.H.
JOURNAL Direct Submission
SUBMITTED (30-DEC-1999) Genome Sequencing Center, Washington
UNIVERSITY School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Mar 15, 2000 this sequence version replaced gi:7139944.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0011A18
----- Summary Statistics -----
Sequencing vector: M13: 77%
Sequencing vector: plasmid: 23%
Chemistry: Dye-terminator Big Dye: 23% of reads
Chemistry: Dye-terminator Big Dye: 23% of reads
Assembly program: Phrap, version 0.990119
Consensus quality: 167434 bases at least Q30

```

```

Consensus quality: 167583 bases at least Q20
Insert size: 154000; agarose-fp
Insert size: 168336; sum-of-contigs
Quality coverage: 8.04 in Q20 bases; agarose-fp
Quality coverage: 7.38 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 4227 4326: contig of 4226 bp in length
* 4327 4326: gap of unknown length
* 14888 14897: contig of 10571 bp in length
* 14898 14997: gap of unknown length
* 53197 53196: contig of 38199 bp in length
* 53297 53296: gap of unknown length
* 97272 97271: contig of 43975 bp in length
* 97372 97371: gap of unknown length
* 97372 168736: contig of 71365 bp in length.
Location/Qualifiers
1. 168736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-11A18"
1. 4226
/note="assembly_name:Contig14"
4327. 14897
/note="assembly_name:Contig15"
14998. 53196
/note="assembly_name:Contig16
clone_end:T7
vector_side:left"
53297. 97271
/note="assembly_name:Contig17"
97372. 168736
/note="assembly_name:Contig18"
BASE COUNT 55308 a 29076 c 28902 g 55038 t 412 others
ORIGIN

```

```

Query Match 1.1%; Score 31; DB 2; Length 168736;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2704 cctgacccctctccactgagagctgtct 2734
Db 100190 CCTGCATCCCTCTCCACTGAGAGCTGTCT 100160

```

```

RESULT 29
AC024996/c 177864 bp DNA linear PRI 12-DEC-2001
LOCUS Homo sapiens chromosome 8, clone RP11-697C18, complete sequence.
DEFINITION AC024996
ACCESSION AC024996
VERSION AC024996.6 GI:17530778
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 177864)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-697C18
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 177864)
REFERENCE Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,

```

Campopiano, A., Castie, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, I., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Plerie, N., Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 177864)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerie, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (29-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 177864)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerie, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (12-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Dec 12, 2001 this sequence version replaced gi:15808595. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L7519
Center clone name: 697_C_18

FEATURES
source

Location/Qualifiers
1. 177864
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-697C18"
/clone.lib="RPC1-11 Human Male BAC"
224..251
repeat_region
/rpt_family="AT_rich"
complement(529..607)
/rpt_family="MBR112"
complement(671..963)
/rpt_family="Alusx"
997..1033
repeat_region
/rpt_family="AT_rich"
1216..1304
/rpt_family="MTR"
1719..1872
/rpt_family="L3"
2006..2195
/rpt_family="MTR"
2325..2361
/rpt_family="AT_rich"
complement(2899..3200)
/rpt_family="Aluub"
3277..3318
/rpt_family="AT_rich"
4619..5073
/rpt_family="L2"
complement(5429..5821)
/rpt_family="L2"
complement(5819..5983)
/rpt_family="MTR"
complement(7391..7637)
/rpt_family="MER46A"
7821..7849
/rpt_family="(T5)n"
complement(7971..8112)
/rpt_family="Charlie4"
complement(8829..9127)
/rpt_family="Alusx"
complement(9280..9434)
/rpt_family="L2"
9767..9963
/rpt_family="L1MB6"
complement(10153..10245)
/rpt_family="7SK"
10879..10906
/rpt_family="AT_rich"
11089..11141
/rpt_family="CT_rich"
complement(11500..11898)
/rpt_family="MS7B"
complement(11899..13102)
/rpt_family="MS7B-int"
complement(13191..13686)
/rpt_family="LTR27"
complement(13687..14103)
/rpt_family="MS7B-int"

```
repeat_region complement(14106..14473)
                    /rpt_family="THE1B"
repeat_region complement(14513..14922)
                    /rpt_family="MSRB"
repeat_region complement(15295..15349)
                    /rpt_family="L2"
repeat_region 15350..15369
                    /rpt_family="(GAGAA)n"
repeat_region 15991..16303
                    /rpt_family="Alusq"
repeat_region complement(16656..16735)
                    /rpt_family="MADE1"
repeat_region 17693..17831
                    /rpt_family="L2"
repeat_region complement(18046..18221)
                    /rpt_family="MER58C"
repeat_region 18594..18614
                    /rpt_family="(A)n"
repeat_region 21544..21690
                    /rpt_family="MIR"
repeat_region 21741..21946
                    /rpt_family="MIR"
repeat_region 22539..22860
                    /rpt_family="MER1B"
repeat_region 23154..23175
                    /rpt_family="(T)n"
repeat_region 23207..23260
                    /rpt_family="MADE1"
repeat_region 23588..23608
                    /rpt_family="AT_rich"
repeat_region 24080..24101
                    /rpt_family="(A)n"
repeat_region complement(24332..24784)
                    /rpt_family="MLT1C"
repeat_region 25422..25452
                    /rpt_family="AT_rich"
repeat_region complement(26347..26643)
```

Query Match 1.1%; Score 31; DB 9; Length 177864;
Best Local Similarity 100.0%; Pred. No. 7 6e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 2704 cctgcacccctccacacagagctgttct 2734
|||||
Db 13356 CCGTCATCCCTCTCCACTGAGAGCTGTCT 13326

RESULT 30
HSM800283 2328 bp mRNA linear PRI 18-FEB-2000
LOCUS HSM800283 2328 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp564F1422 (from clone DKFZp564F1422).
ACCESSION AL049942
VERSION AL049942.1 GI:4884185
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2328)
AUTHORS Mamubut,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferplatz 18a, D-82152
Munich, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGORA (Berlin/Germany) within the CDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp564F1422) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
Location/Qualifiers
```

```
source
1..2328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp564F1422"
/tissue="brain"
/clone_lib="564 (synonym: hfbp2). Vector pAMP1; host
X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..1492
/gene="DKFZp564F1422"
<1..1492
/gene="DKFZp564F1422"
/feature="djb94b14.1, strong similarity to Zn finger
proteins"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAB43216.1"
/db_xref="GI:4884186"
/db_xref="SPRMBL:09Y3Y5"
/translation="RTSGEKPEICVKGCGYTSKSLVHTERTHGERPYECGCR
RENDKSYNKLKHAHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
SNKSHLTHORTHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
KSLVTHORTHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
TFALHORTHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
TIHRTHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
HOKTHSGEKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
RTHSGKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR
IHGERPYECGCRKPEVCGRGVYTNKSYVHVRHISGKPEVPCGCGR"
2290..2295
polya_signal
polya_site
BASE COUNT 674 a 501 c 594 g 559 t
ORIGIN
```

Query Match 1.0%; Score 29; DB 9; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1659 caccagagacacacacagagagagc 1687
|||||
Db 332 CACCAGAGACACACAGAGAGAGC 360

RESULT 31
AK056466 2790 bp mRNA linear PRI 31-OCT-2001
LOCUS AK056466 2790 bp mRNA linear PRI 31-OCT-2001
DEFINITION Homo sapiens cDNA FLJ131904 f1s, clone NT2RP7004352.
ACCESSION AK056466
VERSION AK056466.1 GI:16551878
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens testis carcinoma cell line:NT2 CDNA to mRNA,
clone lib:NT2RP7 clone:NT2RP7004352.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kunagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiya,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2790)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kiserazu, Chiba 292-0812, Japan
(E-mail:genomices@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
```


CDS

```

/lab_host="DH10B-R"
/notes="Vector: pOT87"
134..2389
/codon_start=1
/product="zinc finger protein 337"
/protein_id="AAH21298.1"
/db_xref="GI:18204656"
/translation="MGPOGARROAFLAGDVTDFQTEKWRLLSPARALRYETLEN
YSHLSIGLHSHPELIRLEGEPMGERRRRRPGACAGTAVHEVLRPNLAIOR
OOLOPDSODSFDAEGCEKSTKPMASPPILRAVSRNRYVELESSQORE
NPTEDIKVLKIGINSRNGAKCAERGDGRKMMVLIHKASHRQKLFTRCHQGR
DESAILLHONTHTGKSYVCSVCGRFSKANLLRHQTHSGKRPCKVCGGYTSK
SYLVHETHTGKPYECOCGRFNDKSSYNKHLAHSKSEKPFVCEKCGYTSK
FYVAKRIHSEKPYRCOECGRFSNKLITTHQTHSGKPFACROCKOSFYKSL
RHQTHSGKPFVCKOERFSQKSTLYVHQRHSGKPFVCEKCGYTSKSTLYKH
QTHSEKPFVCKOERFSQKSTLYVHQRHSGKPFVCEKCGYTSKSTLYKH
AHGKERFRCDCGRGFTLKLITLHQRHSGKPFVCEKCGYTSKSTLYKH
SGRPNCKDCGRGFTLKLITLHQRHSGKPFVCEKCGYTSKSTLYKH
KOPFVCEKCGRGNMKNLITLHQRHSGKPFVCKOERFSQKSTLYKH
KPFVCEKCGRGNMKNLITLHQRHSGKPFVCKOERFSQKSTLYKH
FCTGSYGAS"
FCGSYGAS"

```

BASE COUNT 988 a 721 c 844 g 720 t

Query Match 1.0%; Score 29; DB 9; Length 3273;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggaggaagcc 1687
 Db 1229 CACCAGAGACACTCAGGAGGAGACC 1257

RESULT 34
 HS694B14/c
 LOCUS
 DEFINITION Human DNA sequence from clone RP4-694B14 on chromosome 20p11.1-11.22 Contains a novel KRAB box protein with 18 C2H2 type zinc finger domains, a novel haloacid dehalogenase-like hydrolase family protein similar to (archaea) bacterial proteins, two putative novel genes, a novel pseudogene, ESTs, an STS, GSSs and three CPG islands, complete sequence.

ACCESSION AL031673 127867 bp DNA linear PRI 04-APR-2001
 VERSION AL031673.19 GI:11968366
 KEYWORDS HTG; CPG island; haloacid dehalogenase; hydrolase; KRAB box; zinc finger.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 127867)
 AUTHORS Smith, M.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

On Dec 22, 2000 this sequence version replaced gi:10198607.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Bm: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr20>

RP4-694B14 is from the library RPI-4 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP4-694B14. The true left end of clone RPI3-401N8 is at 116161 in this sequence. The true right end of clone RPI1-9616 is at 20586 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES

source	location/Qualifiers
repeat_region	1..127867
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="20"
repeat_region	/map="p11.1-11.22"
repeat_region	/clone="RP4-694B14"
repeat_region	/clone_id="RPI-4"
repeat_region	1..74
repeat_region	/note="Aluub repeat: matches 194..266 of consensus"
repeat_region	787..1088
repeat_region	/note="AluSp repeat: matches 1..303 of consensus"
repeat_region	1513..1786
repeat_region	/note="AluX repeat: matches 1..295 of consensus"
repeat_region	1838..2312
repeat_region	/note="L1MC3 repeat: matches 7230..7739 of consensus"
repeat_region	2373..2669
repeat_region	/note="Aluub repeat: matches 1..298 of consensus"
repeat_region	2670..3427
repeat_region	/note="L1MC3 repeat: matches 6508..7230 of consensus"
repeat_region	3428..3725
repeat_region	/note="Aluub repeat: matches 1..298 of consensus"
repeat_region	3726..4094
repeat_region	/note="L1MC3 repeat: matches 6157..6508 of consensus"
repeat_region	4095..4449
repeat_region	/note="TBE1A repeat: matches 1..354 of consensus"
repeat_region	4450..4592
repeat_region	/note="L1MC3 repeat: matches 6022..6157 of consensus"
repeat_region	4593..4898
repeat_region	/note="AluSc repeat: matches 1..306 of consensus"
repeat_region	4899..5002
repeat_region	/note="L1MC3 repeat: matches 5924..6022 of consensus"
repeat_region	5003..5309
repeat_region	/note="Aluuo repeat: matches 1..302 of consensus"
repeat_region	5310..5576
repeat_region	/note="L1MC3 repeat: matches 5460..5924 of consensus"
repeat_region	5904..6210
repeat_region	/note="AluX repeat: matches 1..307 of consensus"
repeat_region	6481..6579
repeat_region	/note="L2 repeat: matches 2634..2743 of consensus"
repeat_region	6626..7016
repeat_region	/note="MLT1B repeat: matches 6..390 of consensus"
repeat_region	7565..7652
repeat_region	/note="44 copies 2 mer ct 73% conserved"
repeat_region	7668..7781
repeat_region	/note="57 copies 2 mer ct 75% conserved"
repeat_region	complement(7844..8286)
repeat_region	/note="Random repeat: region contains forced join and single clone. HindIII digest suggests there is approximately 1kb missing from the tandem repeat."
repeat_region	8695..8132
repeat_region	/note="19 copies 2 mer ca 78% conserved"
repeat_region	complement(8206..8578)
repeat_region	/note="match: GSS: Em:A0535283"
repeat_region	complement(8211..8572)
repeat_region	/note="match: GSS: Em:A0507005"
repeat_region	complement(8265..8579)
repeat_region	/note="match: GSS: Em:A0587567"
repeat_region	8822..9429

```

/note="cpg_island"
/evidence=not_experimental
10072. .10203
/note="FLAM_C repeat: matches 2. .133 of consensus"
repeat__region
10737. .11028
/note="AluX repeat: matches 6. .298 of consensus"
repeat__region
11364. .11671
/note="AluX repeat: matches 1. .304 of consensus"
repeat__region
12208. .12496
/note="AluX repeat: matches 1. .299 of consensus"
repeat__region
12508. .12813
/note="AluB repeat: matches 1. .301 of consensus"
repeat__region
13209. .13290
/note="MLTIF repeat: matches 243. .317 of consensus"
repeat__region
13319. .13366
/note="MLTIF repeat: matches 280. .327 of consensus"
repeat__region
13394. .13480
/note="MLTIF repeat: matches 468. .547 of consensus"
repeat__region
13501. .13636
/note="FLAM_C repeat: matches 1. .133 of consensus"
repeat__region
13828. .14126
/note="AluB repeat: matches 1. .304 of consensus"
misc__feature
complement(13929. .14330)
/note="match: GSS: Em:AQ375915"
14518. .14584
/note="L2 repeat: matches 2677. .2745 of consensus"
repeat__region
14591. .14757
/note="L2 repeat: matches 1709. .1875 of consensus"
repeat__region
14937. .15009
/note="AluI/FLAM repeat: matches 1. .73 of consensus"
misc__feature
complement(15027. .15313)
/note="match: GSS: Em:AQ280927"
complement(15040. .15332)
/note="match: STS: Em:HS528L19S"
15040. .15332
/note="AluX repeat: matches 1. .293 of consensus"
misc__feature
complement(15061. .15328)
/note="match: GSS: Em:AS18424"
complement(15065. .15333)
/note="match: GSS: Em:AQ746595"
complement(15066. .15246)
/note="match: GSS: Em:AQ807616"
complement(15089. .15356)
/note="match: GSS: Em:AQ490556"
15091. .15333
/note="match: STS: Em:G20204"
complement(15097. .15333)
/note="match: GSS: Em:AS521752"
15109. .15332
/note="match: GSS: Em:B01680"
complement(15111. .15332)
/note="match: GSS: Em:AS521733"
15114. .15332
/note="match: GSS: Em:AQ352998"
15117. .15332
/note="match: GSS: Em:AS519283"
15121. .15331
/note="match: GSS: Em:AS520451"
15121. .15332
/note="match: GSS: Em:AQ223482 Em:AQ631465"
complement(15127. .15356)
/note="match: GSS: Em:AQ115261"
15129. .15340
/note="match: GSS: Em:AQ532262"
complement(15135. .15332)
/note="match: GSS: Em:AQ618206"
15143. .15329
/note="match: GSS: Em:AQ791131"
15156. .15348
/note="match: GSS: Em:AQ536985"
complement(15160. .15332)
/note="match: STS: Em:G42878"
complement(15171. .15346)

```

```

misc_feature /note="match: GSS: Em:AQ245808  
complement(15171..15344)"  
misc_feature /note="match: STS: Em:G15778"  
complement(15182..15328)  
misc_feature /note="match: GSS: Em:AQ315996"  
complement(15188..15332)  
misc_feature /note="match: STS: Em:G07766"  
15192..15332  
misc_feature /note="match: GSS: Em:AQ279850"  
complement(15194..15356)  
misc_feature /note="match: GSS: Em:AQ081971"  
complement(15200..15333)  
misc_feature /note="match: GSS: Em:B88705"  
15207..15331  
misc_feature /note="match: STS: Em:HS198VD10"  
15216..15353  
misc_feature /note="match: GSS: Em:AG014867"  

Query Match 1.0% Score 29: DB 9: Length 127867;  
Best Local Similarity 100.0%: Pred. No. 0.00011;  
Matches 29: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1659 caccagagacacactcagggagaagcc 1687  
|||||  
Db 82604 CACCAGAGACACTCAGGAGAGAGCC 82576
RESULT 35  
AC013770  
LOCUS AC013770  
DEFINITION Homo sapiens clone RP11-3B10, WORKING DRAFT SEQUENCE, 15 unordered  
pieces:  
AC013770  
AC013770.3 GI:7280299  
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 170129)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens, clone RP11-3B10  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 170129)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donejan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,  
Galland,J.C., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lefkocky,J., Lieu,C., Locke,K., Macdonald,P., Marguis,N.,  
McEwan,P., McGurt,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,P., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testafae,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
TITLE Direct Submission  
JOURNAL Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L2702  
Center clone name: 3_B-10

```

```

----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161346 bases at least Q40
Consensus quality: 165569 bases at least Q30
Consensus quality: 167390 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 168729; sum-of-ctnigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1069: contig of 1069 bp in length
* 1070 1169: gap of 100 bp
* 1170 2602: contig of 1433 bp in length
* 2603 2702: gap of 100 bp
* 2703 4232: contig of 1530 bp in length
* 4233 4332: gap of 100 bp
* 4333 5881: contig of 1549 bp in length
* 5882 8954: gap of 100 bp
* 8955 9054: gap of 100 bp
* 9055 11624: contig of 2570 bp in length
* 11625 11724: gap of 100 bp
* 11725 15457: contig of 3733 bp in length
* 15458 15557: gap of 100 bp
* 15558 23562: contig of 8005 bp in length
* 23563 23662: gap of 100 bp
* 23663 31772: contig of 8110 bp in length
* 31773 31872: gap of 100 bp
* 31873 43485: contig of 11613 bp in length
* 43486 43585: gap of 100 bp
* 43586 55357: contig of 11772 bp in length
* 55358 55457: gap of 100 bp
* 55458 66468: contig of 11011 bp in length
* 66469 66568: gap of 100 bp
* 66569 95233: contig of 28665 bp in length
* 95234 95333: gap of 100 bp
* 95334 124230: contig of 28897 bp in length
* 124231 124330: gap of 100 bp
* 124331 170129: contig of 45799 bp in length.

FEATURES
Source
1.170129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="RPCT-11 Human Male BAC"
1.1069
/note="assembly_fragment"
misc_feature
1170.2602
/note="assembly_fragment"
misc_feature
2703.4232
/note="assembly_fragment"
misc_feature
4333.3881
/note="assembly_fragment"
misc_feature
5982.8954
/note="assembly_fragment"
misc_feature
9055.11624
/note="assembly_fragment"
misc_feature
11725.15457
/note="assembly_fragment"
misc_feature
vector_end:T7
vector_side:left"

-----
Query Match 1.0%; Score 29; DB 2; Length 170129;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2470 cctgtctccattgtgtcttctc 2498
Db 33385 CCTGCTTCCCATTTGGTGCTTCCGC 33413
|||||
|||||

RESULT 36
AC109473 181770 bp DNA linear HTG 04-FEB-2002
AC109473/c Homo sapiens chromosome 5 clone RP11-3B10, WORKING DRAFT SEQUENCE,
DEFINITION
9 unordered pieces.
AC109473
AC109473.1 GI:18483469
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
VERSION
human.
KEYWORDS
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 181770)
DOE Joint Genome Institute.
AUTHORS
Sequencing of Human Chromosome 5
TITLE
Unpublished
JOURNAL
2 (bases 1 to 181770)
DOE Joint Genome Institute.
REFERENCE
Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS
Direct Submission
TITLE
Submitted
JOURNAL
COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----Project Information
Center Project Name: 407551
Center clone name: RPCT-11_3B10
-----
Summary Statistics
Consensus quality: 174095 bases at least Q40
Consensus quality: 175269 bases at least Q30
Consensus quality: 176156 bases at least Q20
Estimated insert size: 174280; agarose-fp estimation
Estimated insert size: 180970; sum-of-ctnigs estimation
Quality coverage: 43.3 in Q20 bases; agarose-fp estimation
Quality coverage: 41.7 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```


COMMENT

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'-6' 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES

source

Location/Qualifiers
1..3055
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KIDNE2000051"
/tissue_type="kidney"
/clone_1lb="KIDNE2"
/note="cloning vector: pME18SFL3"
188..1717
/note="unnamed protein product"
/codon_start=-1
/protein_id="BA071000.1"
/db_xref="gi:16550537"
/translation="MLENPNILSLAEPKPEITYTSCSLAFSCQFLSHQVIFIG
LCAENHPGNSSPGHKKQGOQYSHVSCPFENAEGERGGSPMSARTREETSIA
FSPFLQKASAPRKNVETEPSSAQRPNPVDLKGLETLREGLAICREYEDP
NLSNFTINRPTLLGKRPYICSDGSRFSDKSTLIRHHRTHSEKPYVCEGSGFESQ
KSNLSRTHRTSEKPYLRCGSGFESKSLINRQWTHSEKPYVCEGSGFESQ
SFIRHRTSEKPYVCEGSGFESKSLIRHRTSEKPYVCEGSGFESQ
IKHRTHDEKPYVCEGSGFESKSLIRHRTSEKPYVCEGSGFESQ
HRTHSEKPYVCEGSGFESKSLIRHRTSEKPYVCEGSGFESQ
RHSGEKPYVCEGSGFESKSLIRHRTSEKPYVCEGSGFESQ
H"

CDS

BASE COUNT 866 a 692 c 808 g 689 t
ORIGIN

Query Match 1.0%; Score 28; DB 9; Length 3055;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1659 caccagagacacactcaggaggaagc 1686
|||||
Db 1364 CACCAGAGCAGACACTCAGGGAGAGC 1391

RESULT 39
AC091936 61452 bp DNA linear HTG 09-JUN-2001
LOCUS
DEFINITION Homo sapiens chromosome 5 clone RP11-325C20, WORKING DRAFT
SEQUENCE 16 unordered pieces.
AC091936
AC091936.1 GI:14333872
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 61452)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 61452)
DOE Joint Genome Institute.
Direct Submission
Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 531233

Center clone name: RP11-325C20

Summary Statistics

Consensus quality: 47683 bases at least Q40
Consensus quality: 53520 bases at least Q30
Consensus quality: 56595 bases at least Q20
Estimated insert size: 160000; agarose-fp estimation
Estimated insert size: 59952; sum-of-contigs estimation
Quality coverage: 5.17 in Q20 bases; agarose-fp estimation
Quality coverage: 13.8 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1158: contig of 1158 bp in length
* 1159 1258: gap of unknown length
* 1259 2621: contig of 1363 bp in length
* 2622 2721: gap of unknown length
* 2722 3952: contig of 1231 bp in length
* 3953 4052: gap of unknown length
* 4053 5251: contig of 1199 bp in length
* 5252 5351: gap of unknown length
* 5352 6514: contig of 1163 bp in length
* 6515 6615: gap of unknown length
* 6615 7791: contig of 1177 bp in length
* 7792 7891: gap of unknown length
* 7892 9549: contig of 1658 bp in length
* 9550 9650: gap of unknown length
* 9650 11835: contig of 2186 bp in length
* 11836 13920: gap of unknown length
* 13921 14020: contig of 1985 bp in length
* 14021 15094: gap of unknown length
* 15095 15194: contig of 1074 bp in length
* 15195 16490: gap of unknown length
* 16491 16590: contig of 1296 bp in length
* 16591 19279: gap of unknown length
* 19280 19379: contig of 2689 bp in length
* 19380 29586: gap of unknown length
* 29587 29686: gap of unknown length
* 29687 39306: contig of 10207 bp in length
* 39307 39406: gap of unknown length
* 39407 39406: gap of unknown length
* 39407 48478: contig of 9072 bp in length
* 48479 48578: gap of unknown length
* 48579 61452: contig of 12874 bp in length.
Location/Qualifiers
1..61452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-325C20"
/clone_1lb="RP11-325C20"
/clone_1lb="RP11-325C20"

FEATURES

source

BASE COUNT 18087 a 10746 c 10505 g 20606 t 1508 others
ORIGIN

Query Match 1.0%; Score 28; DB 2; Length 61452;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2501 atgattcccaacccctcattattac 2528
|||||
Db 22636 ATGATCCCAACCCCTCATTATTAC 22663

RESULT 40
AF161800 62002 bp DNA linear HTG 13-JUN-2001
LOCUS
DEFINITION Homo sapiens chromosome 8 clone XX-189m5 map 8q21.2, *** SEQUENCING

IN PROGRESS ***, 1 ordered pieces.

AF161800
 VERSION GI:5353769
 HTG: HTGS_PHASE2: HTGS_FUULLTOP.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 62002)
 Schilbhel,M.B., Baumgart,C., Blechschmidt,K., Dete,M., Jahn,N., Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A., Sidiq,R., Taudien,S., Wen,G., Korenberg,J., Rosenthal,A. and Platzer,M.

TITLE Chromosome 8 genomic sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 62002)
 AUTHORS Schudy,A., Schilbhel,M., Schlegelberger,B., Drescher,B., Weber,J., Schattevoy,R., Menzel,U. and Rosenthal,A.

TITLE Direct Submission
 JOURNAL Submitted (23-JUN-1999) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

COMMENT * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 62002: contig of 62002 bp in length.
 Location/Qualifiers
 source 1. 62002
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q21.2"
 /clone="XX-189m5"
 BASE COUNT 17935 a 11598 c 12429 g 20038 t 2 others
 ORIGIN

Query Match 1.0%; Score 28; DB 2; Length 62002;
 Best Local Similarity 100.0%; Pred. No. 0.00045;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 tgcatagcagagcagcagcgaatgcc 2385
 |||||||||||||||||||||||||||||
 Db 5666 TGATAGTGCAGGAGGACATCAATGCC 5639

RESULT 41
 AC105030 65236 bp DNA linear HTG 22-DEC-2001
 LOCUS Homo sapiens chromosome 17 clone CTD-2244F11 map 17, LOW-PASS
 DEFINITION
 AC105030
 VERSION
 AC105030.1 GI:17977409
 HTG: HTGS_PHASE0.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 65236)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 17, clone CTD-2244F11
 Unpublished
 2 (bases 1 to 65236)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhgalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,U.S., Dodge,S., Faro,S.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelis,C., Larroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPeckers,R., Meldrum,J., Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT * NOTE: This record contains 80 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 720: contig of 720 bp in length
 * 721 820: gap of 100 bp
 * 821 1547: contig of 727 bp in length
 * 1548 1647: gap of 100 bp
 * 1648 2393: contig of 746 bp in length
 * 2394 2493: gap of 100 bp
 * 2494 3191: contig of 698 bp in length
 * 3192 3291: gap of 100 bp
 * 3292 3992: contig of 701 bp in length
 * 3993 4092: gap of 100 bp
 * 4093 4779: contig of 687 bp in length
 * 4780 4879: gap of 100 bp
 * 4880 5581: contig of 702 bp in length
 * 5582 5681: gap of 100 bp
 * 5682 6408: contig of 727 bp in length
 * 6409 6508: gap of 100 bp
 * 6509 7240: contig of 732 bp in length
 * 7241 7340: gap of 100 bp
 * 7341 8033: contig of 693 bp in length
 * 8034 8133: gap of 100 bp
 * 8134 8660: contig of 727 bp in length
 * 8661 8960: gap of 100 bp
 * 8961 9698: contig of 738 bp in length
 * 9699 9798: gap of 100 bp
 * 9799 10541: contig of 743 bp in length
 * 10542 10641: gap of 100 bp
 * 10642 11366: contig of 725 bp in length
 * 11367 11466: gap of 100 bp
 * 11467 12162: contig of 696 bp in length
 * 12163 12262: gap of 100 bp

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L23011
 Center clone name: 2244_F_11

```

* 1263 12967: contig of 705 bp in length
* 12968 13067: gap of 100 bp
* 13068 13808: contig of 741 bp in length
* 13809 13908: gap of 100 bp
* 13909 14621: contig of 713 bp in length
* 14622 14721: gap of 100 bp
* 14722 15434: contig of 713 bp in length
* 15435 15534: gap of 100 bp
* 15535 16257: contig of 723 bp in length
* 16258 16357: gap of 100 bp
* 16358 17074: contig of 717 bp in length
* 17075 17174: gap of 100 bp
* 17175 17901: contig of 727 bp in length
* 17902 18001: gap of 100 bp
* 18002 18716: contig of 715 bp in length
* 18717 18816: gap of 100 bp
* 18817 19523: contig of 707 bp in length
* 19524 19623: gap of 100 bp
* 19624 20307: contig of 684 bp in length
* 20308 20407: gap of 100 bp
* 20408 21109: contig of 702 bp in length
* 21110 21209: gap of 100 bp
* 21210 21913: contig of 704 bp in length
* 21914 22013: gap of 100 bp
* 22014 22735: contig of 722 bp in length
* 22736 22835: gap of 100 bp
* 22836 23556: contig of 721 bp in length
* 23557 23656: gap of 100 bp
* 23657 24379: contig of 723 bp in length
* 24380 24479: gap of 100 bp
* 24480 25134: contig of 655 bp in length
* 25135 25234: gap of 100 bp
* 25235 25962: contig of 728 bp in length
* 25963 26062: gap of 100 bp
* 26063 26803: contig of 741 bp in length
* 26804 26903: gap of 100 bp
* 26904 27638: contig of 735 bp in length
* 27639 27738: gap of 100 bp
* 27739 28427: contig of 689 bp in length
* 28428 28527: gap of 100 bp
* 28528 29224: contig of 697 bp in length
* 29225 29324: gap of 100 bp
* 29325 30033: contig of 709 bp in length
* 30034 30133: gap of 100 bp
* 30134 30836: contig of 703 bp in length
* 30837 30936: gap of 100 bp
* 30937 31636: contig of 700 bp in length
* 31637 31736: gap of 100 bp
* 31737 32481: contig of 745 bp in length
* 32482 32581: gap of 100 bp
* 32582 33287: contig of 706 bp in length
* 33288 33387: gap of 100 bp
* 33388 34096: contig of 709 bp in length
* 34097 34196: gap of 100 bp
* 34197 34915: contig of 719 bp in length
* 34916 35015: gap of 100 bp
* 35016 35737: contig of 722 bp in length
* 35738 35837: gap of 100 bp
* 35838 36542: contig of 705 bp in length
* 36543 36642: gap of 100 bp
* 36643 37357: contig of 715 bp in length
* 37358 37457: gap of 100 bp
* 37458 38146: contig of 689 bp in length
* 38147 38246: gap of 100 bp
* 38247 38966: contig of 720 bp in length
* 38967 39066: gap of 100 bp
* 39067 39788: contig of 722 bp in length
* 39789 39888: gap of 100 bp
* 39889 40623: contig of 735 bp in length
* 40624 40723: gap of 100 bp
* 40724 41446: contig of 723 bp in length
* 41447 41546: gap of 100 bp
* 41547 42259: contig of 713 bp in length

```

```

* 42260 42359: gap of 100 bp
* 42360 43092: contig of 733 bp in length
* 43093 43192: gap of 100 bp
* 43193 43918: contig of 726 bp in length
* 43919 44018: gap of 100 bp
* 44019 44746: contig of 728 bp in length
* 44747 44846: gap of 100 bp
* 44847 45554: contig of 708 bp in length
* 45555 45654: gap of 100 bp
* 45655 46368: contig of 714 bp in length
* 46369 46468: gap of 100 bp
* 46469 47197: contig of 729 bp in length
* 47198 47297: gap of 100 bp
* 47298 48020: contig of 723 bp in length
* 48021 48120: gap of 100 bp
* 48121 48851: contig of 731 bp in length
* 48852 48951: gap of 100 bp
* 48952 49667: contig of 716 bp in length
* 49668 49767: gap of 100 bp
* 49768 50502: contig of 735 bp in length
* 50503 50602: gap of 100 bp
* 50603 51289: contig of 687 bp in length
* 51290 51389: gap of 100 bp
* 51390 52124: contig of 735 bp in length
* 52125 52224: gap of 100 bp
* 52225 52885: contig of 661 bp in length
* 52886 52985: gap of 100 bp
* 52986 53698: contig of 713 bp in length
* 53699 53798: gap of 100 bp
* 53799 54520: contig of 722 bp in length
* 54521 54620: gap of 100 bp
* 54621 55337: contig of 717 bp in length
* 55338 55437: gap of 100 bp
* 55438 56170: contig of 733 bp in length

```

```

Query Match      1.0%: Score 28; DB 2; Length 65236;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2461 aattgaagacctgtccatcttggtg 2488
Db 3609 aattgagacacctgtcttccatttgctg 3636

```

```

RESULT 42
AL512489
LOCUS
DEFINITION Human DNA sequence from clone RP11-44708 on chromosome 6, complete
sequence.
ACCESSION AL512489
VERSION AL512489.11 GI:13992076
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
1 (bases 1 to 80846)
REFERENCE
AUTHORS requests: clonerequest@sanger.ac.uk
TITLE Submitted (04-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
request: clonerequest@sanger.ac.uk
COMMENT
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such

```

as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM, EMBL, SW, SSISSPNOT, Tr, TREMBL, Wp, WORMEP. Information on the WORMEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> Rpl1-44708 is from the library RPl-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RPl1-44708 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RPl-162C6 is at 80747 in this sequence. The true right end of clone RPl-271G9 is at 100 in this sequence.

FEATURES

SOURCE

```
1. 80846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Rpl1-44708"
/clone.lib="RPl-11.2"
1..268
/note="L1P repeat: matches 3697. .3964 of consensus"
repeat_region
380..448
/note="3 copies 23 mer 82% conserved"
repeat_region
430..473
/note="11 copies 4 mer tagg 86% conserved"
repeat_region
506..595
/note="45 copies 2 mer ta 62% conserved"
repeat_region
508..591
/note="21 copies 4 mer taga 84% conserved"
repeat_region
2590..2623
/note="17 copies 2 mer tg 85% conserved"
repeat_region
3664..3969
/note="L1MB8 repeat: matches 5847. .6173 of consensus"
repeat_region
4687..4765
/note="L2 repeat: matches 2098. .2176 of consensus"
repeat_region
5352..5419
/note="34 copies 2 mer at 75% conserved"
repeat_region
6233..6631
/note="Charliela repeat: matches 1025. .1455 of consensus"
repeat_region
6632..7214
/note="L1R28 repeat: matches 1. .608 of consensus"
repeat_region
7279..7323
/note="MER52A repeat: matches 1207. .1252 of consensus"
repeat_region
7361..7515
/note="MER52A repeat: matches 1602. .1750 of consensus"
repeat_region
7543..8513
/note="Charliela repeat: matches 6. .1022 of consensus"
repeat_region
8659..8961
/note="Alusq repeat: matches 1. .302 of consensus"
misc_feature
complement(9091..9549)
/note="match: GSS: Em:AQ728035"
repeat_region
9294..9423
/note="FLAM_C repeat: matches 1. .131 of consensus"
misc_feature
9582..10277
/note="match: GSS: Em:AQ626592"
repeat_region
11212..11251
/note="20 copies 2 mer tt 87% conserved"
repeat_region
12790..12897
/note="MIR repeat: matches 118. .232 of consensus"
misc_feature
complement(12908..13388)
/note="match: GSS: Em:AQ437578"
repeat_region
13359..13420
/note="MIR repeat: matches 17. .78 of consensus"
misc_feature
13377..14279
```

```
repeat_region
14124..14355
/note="match: GSS: Em:AQ748923"
repeat_region
14357..14562
/note="L2 repeat: matches 2514. .2748 of consensus"
repeat_region
15005..15315
/note="L1MA2 repeat: matches 6087. .6308 of consensus"
repeat_region
15710..15953
/note="AluY repeat: matches 1. .308 of consensus"
repeat_region
16730..17022
/note="MIR repeat: matches 5. .262 of consensus"
misc_feature
complement(17405..17893)
/note="match: GSS: Em:AQ672286"
repeat_region
18033..18161
/note="MIR repeat: matches 23. .153 of consensus"
repeat_region
20356..20542
/note="MIR repeat: matches 2. .189 of consensus"
repeat_region
20701..20897
/note="MER63A repeat: matches 1. .200 of consensus"
repeat_region
21189..21273
/note="L2 repeat: matches 2624. .2710 of consensus"
repeat_region
21514..21663
/note="MIR repeat: matches 77. .231 of consensus"
repeat_region
21797..22062
/note="L1PA6 repeat: matches 5233. .5497 of consensus"
repeat_region
22063..22700
/note="L1PA6 repeat: matches 5497. .6141 of consensus"
misc_feature
23497..23826
/note="match: GSS: Em:B80555"
misc_feature
23505..23963
/note="match: GSS: Em:AQ235419"
repeat_region
23934..23959
/note="L3 copies 2 mer aa 100% conserved"
repeat_region
25593..31056
/note="L1PB1 repeat: matches -279. .6155 of consensus"
repeat_region
31044..32076
/note="L1PB1 repeat: matches -1537. -.412 of consensus"
repeat_region
32085..32551
/note="Tiger2a repeat: matches 2. .434 of consensus"
repeat_region
32852..32907
/note="28 copies 2 mer tg 76% conserved"
repeat_region
32872..32907
/note="9 copies 4 mer tctg 91% conserved"
misc_feature
complement(33777..34116)
/note="match: GSS: Em:A2246003"
misc_feature
34248..34564
/note="match: GSS: Em:AQ791133"
misc_feature
complement(35522..36045)
/note="match: GSS: Em:AQ561011"
repeat_region
35729..41433
/note="L1PA7 repeat: matches 549. .6140 of consensus"
repeat_region
41074..41617
/note="match: GSS: Em:AQ231043"
repeat_region
42001..42188
/note="L2 repeat: matches 2548. .2741 of consensus"
repeat_region
42244..42338
/note="MER5A repeat: matches 13. .105 of consensus"
repeat_region
42374..42653
/note="L2 repeat: matches 2246. .2519 of consensus"
repeat_region
43032..43265
/note="MIR repeat: matches 13. .262 of consensus"
repeat_region
43855..44065
/note="M1T1J repeat: matches 293. .516 of consensus"
repeat_region
44137..44860
/note="L1PA7 repeat: matches 5405. .6141 of consensus"
repeat_region
44930..44967
/note="19 copies 2 mer ta 78% conserved"
repeat_region
45003..45194
/note="96 copies 2 mer at 58% conserved"
repeat_region
45135..45194
/note="15 copies 4 mer atat 75% conserved"
repeat_region
45200..45237
/note="19 copies 2 mer at 78% conserved"
```

```
repeat_region 46224..46423 /note="AluSg/x repeat: matches 113. .312 of consensus"
repeat_region 46477..46617 /note="HERVL repeat: matches 3164. .3304 of consensus"
repeat_region 47132..47713 /note="L1MEC repeat: matches 1392. .1977 of consensus"
repeat_region 47950..47991 /note="L1MEC repeat: matches 2272. .2313 of consensus"
repeat_region 48550..48762 /note="L1ME1 repeat: matches 5956. .6168 of consensus"
misc_feature 50034..50527 /note="match: GSS: Em:AQ732908"
repeat_region 50281..51058 /note="L1PA2 repeat: matches 4. .776 of consensus"
repeat_region 51054..56274 /note="L1PA2 repeat: matches 900. .6146 of consensus"
repeat_region 56466..56574 /note="AluSg/x repeat: matches 1. .109 of consensus"
repeat_region 56577..56785 /note="MLT2E repeat: matches 117. .337 of consensus"
repeat_region 57436..57706 /note="AluSg repeat: matches 5. .287 of consensus"
repeat_region 59298..60448 /note="L1ME repeat: matches 890. .1593 of consensus"
repeat_region 60497..60587 /note="L1ME repeat: matches 2128. .2214 of consensus"
repeat_region 60588..60904
```

```
Query Match 1.0%: Score 28; DB 9; Length 80846;
Best Local Similarity 100.0%: Pred. No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2358 tgatagtcgacgagcagcgaatgcc 2385
Db 6632 TGATAGTCGACGAGCGATCAATGCC 6659
```

```
RESULT 43
HSJ734P14 90220 bp DNA linear PRI 23-FEB-2001
LOCUS Human DNA sequence from clone RP4-734P14 on chromosome 20. Contains
DEFINITION the 3' end of a novel KRAB box protein gene, a novel C/EBP2 zinc
finger domain protein gene, the SNRFB gene for small nuclear
ribonucleoprotein polypeptides B and B1 (B') and the 3' end of the
gene for a novel transglutaminase. Contains ESTs, STSs, GSSs and a
CpG island, complete sequence.
```

```
ACCESSION AL049650
VERSION AL049650.8 GI:5123801
KEYWORDS HTG; CpG island; KRAB box; snRNP; SNRFB; transglutaminase; Zinc
finger.
```

```
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 90220)
Blakey, S.
Direct Submission
Submitted (22-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
```

COMMENT

On Jun 22, 1999 this sequence version replaced gi:4995635. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20> RP4-734P14 is from the library RP4-1 constructed by the group of Pletter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP4-734P14. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP4-734P14 is at 1 in this sequence. The true left end of clone RP4-816K17 is at 90117 in this sequence.

FEATURES

source

repeat_region

mRNA

gene

CDS

```
join(929..1046,1247..1305,1999..2125,2740..2792)
/note="protein homologs suggest that this gene and gene
.5 are part of the same gene in the form of a yet to be
determined isoform
match: proteins: Sw:P52736 Sw:Q06730 Sw:P51523 Sw:Q02386
Sw:P51786 Sw:Q99676 Tr:Q06792 Sw:Q03923 Tr:Q9Y473
Tr:Q9YK04 Tr:Q9P004 Sw:Q95201 Tr:Q9JL03 Sw:Q62396
Tr:Q922X6 Tr:Q15322 Tr:Q9WV10 Tr:Q14913 Sw:Q61751
Tr:Q9P0L1 Tr:Q9NS42 Sw:Q9UJ03 Tr:P70591 Tr:Q14586
Sw:Q02975 Tr:Q94892 Tr:Q9NSD1 Tr:Q62977 Tr:Q9Y2N8
Tr:Q9Y2N9 Tr:Q43296 Tr:Q9WY4 Tr:Q9Y3W9 Tr:Q9NR11
Sw:Q60765 Tr:Q95765 Tr:Q90X19 Tr:Q9XSRI"
/codon_start=1
/evidence=not_experimental
/product="dJ734P14.1 (novel KRAB box protein)"
/protein_id="CAH46713.2"
/db_xref="GI:13160044"
/translacion="MMLPTPSALGDIYVEITLLPKNGEVEVETMKLLTONHKAKGPSN
DIDCDQKKEGKAQIVPVTFRDVTVIFTEAEKRLSPEDGRNLKYKEMLENYRNLLSLG
QEETILLANIVKSHLY"
1153..1180
/note="MER70B repeat: matches 267. .294 of consensus"
2150..2632
/note="L2 repeat: matches 2288. .2749 of consensus"
2735..2975
/note="AluSg repeat: matches 63. .303 of consensus"
3026..3345
/note="MER58B repeat: matches 2. .341 of consensus"
3356..3684
/note="MER33 repeat: matches -16. .324 of consensus"
4121..4296
/note="AluSg repeat: matches 135. .310 of consensus"
4822..4952
/note="RIAM_C repeat: matches 1. .133 of consensus"
5086..5435
/note="MLT1I repeat: matches 28. .396 of consensus"
```



```

Center/clone name: GSI-225N9
----- Summary Statistics -----
Sequencing vector: M13, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 105932 bases at least Q40
Consensus quality: 107474 bases at least Q30
Consensus quality: 108376 bases at least Q20
Quality coverage: 7.28 x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality 10.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

```

```

FEATURES
source      Location/Qualifiers
1..109217
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"

```

BASE COUNT	32201	a	21918	c	22227	g	32871	t
ORIGIN								

```

Query Match      1.0%; Score 28; DB 9; Length 109217;
Best Local Similarity 100.0%; Pred. NO. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy . 2358 tgaatgtygcaggaagcagtcgaatgcc 2385
      |||||
Db 54651 TGATAGTGGCAGGAGGCAGTCAATGCC 54624

```

RESULT	45
AC099798	
LOCUS	131328 bp DNA linear PRI 03-FEB-2002
DEFINITION	Homo sapiens chromosome 7 clone RP11-559E21, complete sequence.
ACCESSION	AC099798
VERSION	AC099798..4 GI:18482338
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 131328)
Waterston, R. H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 131328)
Waterston, R. H.
Direct Submission
Submitted (21-NOV-2001) Genome Sequencing Center, Washington

REFERENCE 3 (bases 1 to 131328)
AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2002) Genome Sequencing Center, Washington

COMMENT On Feb 3, 2002 this sequence version replaced gi:18056737.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu
----- Project Information
Center project name: H_NH0559E21

```
FEATURES                                location/Qualifiers
source                                  1. 131328
                                        /organism="Homo sapiens"
                                        /db_xref="taxon:9606"
                                        /chromosome="7"
                                        /clone="Rp11-559E21"
BASE COUNT                             35939 a 25310 g 39949 t
ORIGIN
```

Query Match	1.0%	Score 28:	DB 9:	Length 131328:
Best Local Similarity	100.0%	Pred. No.	0.00044:	
Best Match	28:	Conservative	0:	Mismatches 0:
				Indels 0:
				Gaps 0:

OY	2358	tgtatagttgycaggaaggcagtccaattgcc	2385
		t t t t t t t	
Db	130808	TGATAGTGGCAGGAGGCAGTCAATGCC	130835

RESULT	46					
AC011059						
LOCUS	AC011059	143799 bp	DNA	linear	PRI 10-DEC-2001	
DEFINITION	Homo sapiens chromosome 18, clone RP11-11F23, complete sequence.					
ACCSSION	AC011059					
VERSION	AC011059.8	GI:17432875				
KEYWORDS	HTc.					
SOURCE	human.					
ORGANISM	Homo sapiens					

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 143799)	Blirren,B., Linton,L., Nusbaum,C. and Lander,E.	Homo sapiens chromosome 18, clone RP11-11F23
2 (bases 1 to 143799)	Blirren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.	Unpublished

COOKE, P., DERELLANO, K., DEWAR, K., DOMINO, M., DONEGAN, L., DOYLE, M.,
FERREIRA, P., FITZHUUGH, W., FORREST, C., FUNKE, R., GAGE, D.,
GALLAGHER, J., GARDYNA, S., GRANT, G., HAGOS, B., HEAFORD, A., HORTON, L.,
HOWLAND, J. C., JOHNSON, R., JONES, C., KAMM, L., KARLATS, A., KLEIN, J.,
LEBOCKZY, J., LIEN, C., LOCKE, K., MACDONALD, P., MARCUS, N.,
MEWAN, P., MCGURK, A., MCKERNAN, K., MCNAUGHTIN, J., MELDRIM, J.,
MORROW, J., NAYLOR, J., NORMAN, C. H., O'CONNOR, T., O'DONNELL, P.,
PETERSON, K., POLLARA, V., RILEY, R., ROY, A., SANTOS, R., SEEVERY, P.,
STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., TALAMAS, J.,
TESSIERE, S., TITRELL, A., VASSILIER, H., VO, A., WHEELER, J., WU, X.,
WYMAN, D., YE, W. J., ZIMMER, A. and ZODY, M.

Direct Submission

TITLE Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 143799)

Barron, B., Antonin, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barin, N., Bastien, V., Bougslavsky, L., Boukgalter, B., Brown, A.,
 Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
 Collins, S., Collimore, A., Cooke, P., Dearellano, K., Dewar, K.,
 Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,
 Galaçan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W.,
 Illey, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K.,
 Lamaçars, R., Landers, T., Lehocsky, N., Levine, R., Liu, G.,
 Maclean, C., MacDonald, P., Marquis, N., Mathews, C., McCarthy, M.,
 McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneses, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, C., Nguyen, C., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rossetti, M., Roy, A., Santos, R., Schauer, S., Schnuppach, R., Seaman, S.,

Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (31-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 143799)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Choepel, V., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (02-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

5 (bases 1 to 143799)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Alt, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B., Choepel, V., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (10-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Dec 10, 2001 this sequence version replaced gi:14269782.

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L3150
Center clone name: 11_F_23

Location/Qualifiers

source
1. 143799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-11F23"
/clone_lib="RP11 Human Male BAC"
744 . 925
repeat_region
1083 . 1140
/rpt_family="L3"
repeat_region
1586 . 1814
/rpt_family="AluSg"
complement(1912 . 3816)
/rpt_family="L1McC"
complement(4170 . 4300)
/rpt_family="L1McC"
complement(5546 . 5853)
/rpt_family="AluSg"
6814 . 7113
/rpt_family="AluY"
complement(7304 . 7585)
/rpt_family="L1PA7"
complement(9689 . 9812)
/rpt_family="L3"
9963 . 10092
/rpt_family="AluY"
complement(10936 . 11160)
/rpt_family="L3"
11196 . 11263
/rpt_family="MIR"
11292 . 11585
/rpt_family="AluB"
complement(11795 . 12121)
/rpt_family="AluSg"
complement(12152 . 12483)
/rpt_family="MER7A"
12923 . 13014
/rpt_family="MER31A"
12995 . 13140
/rpt_family="Tigger3(Golem)"
13146 . 13416
/rpt_family="MER31A"
complement(13584 . 13922)
/rpt_family="MER1B"
14939 . 14970
/rpt_family="(7)n"
complement(16283 . 16380)
/rpt_family="MIR3"
16440 . 16707
/rpt_family="AluX"
17310 . 17464
/rpt_family="MIR"
complement(17753 . 17830)
/rpt_family="MIR"
18533 . 18717
/rpt_family="MIR3"
18729 . 18916
/rpt_family="MIR3"
complement(18923 . 19163)
/rpt_family="L2"
complement(20725 . 20764)
/rpt_family="L2"
complement(20850 . 20988)
/rpt_family="ETAM.C"
complement(21135 . 21763)
/rpt_family="L2"
22218 . 22261
/rpt_family="CF-rich"
complement(22264 . 22529)
/rpt_family="AluX"

FEATURES

repeat_region 2342. .23396
 /rpt_family="L2"
 repeat_region 23628. .23700
 /rpt_family="L2"
 repeat_region 25505. .25550
 /rpt_family="(TC)n"

Query Match 1.0%; Score 28; DB 9; Length 143799;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 tgatagtgagagagagcgaatgc 2385
 Db 123323 TGATAGTGAGAGAGCAATGCC 123350

RESULT 47
 AC010998 146831 bp DNA linear PRI 09-AUG-2001
 DEFINITION Homo sapiens chromosome 10 clone RP11-95116, complete sequence.
 AC010998
 VERSION AC010998.13 GI:15144282
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 146831)
 Smith,D.R.

TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
 AUTHORS
 JOURNAL
 Unpublished
 Sequence Data

REFERENCE 2 (bases 1 to 146831)
 Smith,D.R.

TITLE Direct Submission
 JOURNAL Submitted (29-SEP-1999) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 3 (bases 1 to 146831)

REFERENCE Smith,D.R.

TITLE Direct Submission

JOURNAL Submitted (09-AUG-2001) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA
 On Aug 9, 2001 this sequence version replaced gi:14717312.

COMMENT Location/Qualifiers

FEATURES
 source
 1. 146831
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-95116"
 /clone_1lb="RP11-95116"
 /clone_1lb="RP11-95116"
 BASE COUNT 41832 a 31937 c 31285 g 41777 t
 ORIGIN

Query Match 1.0%; Score 28; DB 9; Length 146831;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2511 accctacactattacgatactgc 2538
 Db 119994 ACCCTTACCTATTACGTACCTGC 120021

RESULT 48
 AC021765/c 152393 bp DNA linear HTG 20-SEP-2000
 DEFINITION Homo sapiens clone RP11-9D19, WORKING DRAFT SEQUENCE, 11 unordered
 pieces.
 AC021765
 VERSION AC021765.3 GI:10198414
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Unpublished
 2 (bases 1 to 152393)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerley,R., Bede,F.,
 Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
 Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karas,A., Klein,T.,
 Landers,T., Lehoczek,J., Levine,R., Lien,C., Liu,G., Locke,K.,
 Macdonald,P., Marczyk,N., McEwan,P., McGurk,A., McInerney,K.,
 McInerney,R., Meldrum,D., Menus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
 Pierre,N., Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT
 Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 20, 2000 this sequence version replaced gi:1923795.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

Project Information
 Center project name: L3002
 Center clone name: 9_D_19

Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 146933 bases at least Q40
 Consensus quality: 149520 bases at least Q30
 Consensus quality: 150601 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 151393; sum-of-contigs
 Quality coverage: 4.6 in Q20 bases; agarose-fp
 Quality coverage: 4.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 28910: contig of 28910 bp in length
 28911 29010: gap of 100 bp
 29011 31437: contig of 2427 bp in length
 31438 31537: gap of 100 bp
 31538 33952: contig of 2415 bp in length
 33953 34052: gap of 100 bp
 34053 38372: contig of 4320 bp in length
 38373 38472: gap of 100 bp
 38473 81141: contig of 42669 bp in length
 81142 81241: gap of 100 bp
 81242 86894: contig of 5653 bp in length
 86895 86994: gap of 100 bp
 86995 95082: contig of 8088 bp in length
 95083 95182: gap of 100 bp

* 95183 110046: contig of 14864 bp in length
 * 110047 110146: gap of 100 bp
 * 110147 129817: contig of 19671 bp in length
 * 129818 129917: gap of 100 bp
 * 129918 148772: contig of 18855 bp in length
 * 148773 148872: gap of 100 bp
 * 148873 152393: contig of 3521 bp in length.
 Location/Qualifiers
 1. 152393

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-9D19"
 /clone_lib="RP11-11 Human Male BAC"

misc_feature

1. 28910
/note="assembly-fragment"

clone_end:SP6
vector_side:left"

misc_feature

29011..31437
/note="assembly-fragment"

misc_feature

31538..33952
/note="assembly-fragment"

misc_feature

34053..38372
/note="assembly-fragment"

misc_feature

38473..81141
/note="assembly-fragment"

misc_feature

81242..86894
/note="assembly-fragment"

misc_feature

86995..95082
/note="assembly-fragment"

misc_feature

95183..110046
/note="assembly-fragment"

misc_feature

110147..129817
/note="assembly-fragment"

misc_feature

129918..148772
/note="assembly-fragment"

misc_feature

148873..152393
/note="assembly-fragment"

BASE COUNT

49730 a 27512 c 26497 g 47643 t 1011 others

ORIGIN

Query Match 1.0%; Score 28; DB 2; Length 152393;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 tgatagtgcagagagcagcgaatgcc 2385

Db 27507 TGATAGTGCAGCAGCAGCAATGCC 27480

RESULT 49

AC104974/c

LOCUS

DEFINITION

AC104974

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC104974 154604 bp DNA linear HTG 13-FEB-2002
 Homo sapiens chromosome 17 clone RP11-145L16 map 17, WORKING DRAFT
 SEQUENCE, 23 ordered pieces.
 AC104974 GI:18653683
 HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 154604)
 Homo sapiens chromosome 17, clone RP11-145L16
 Unpublished
 2 (bases 1 to 154604)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhgarter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,Y., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

TITLE

JOURNAL

COMMENT

Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hages,B., Hearford,A., Horton,L., Hulme,W., Iller,I., Johnson,R.,
 Jones,C., Kamet,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
 Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Rella,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stenge-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 13, 2002 this sequence version replaced gi:18464184.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L22871

Center clone name: 145_L16

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 148031 bases at least Q40

Consensus quality: 150405 bases at least Q30

Consensus quality: 151462 bases at least Q20

Insert size: 146000; agarose-fp

Insert size: 152404; sum-of-ctdigs

Quality coverage: 9.3 in Q20 bases; agarose-fp

Quality coverage: 8.9 in Q20 bases; sum-of-ctdigs

NOTE: This is a 'working draft' sequence. It currently

consists of 23 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and

the accession number will be preserved.

1 1709: contig of 1709 bp in length

1710 1809: gap of 100 bp

1810 1999: contig of 190 bp in length

2000 2099: gap of 100 bp

2100 3127: contig of 1028 bp in length

3128 3227: gap of 100 bp

3228 4561: contig of 1334 bp in length

4562 4661: gap of 100 bp

4662 5811: contig of 1150 bp in length

5812 5911: gap of 100 bp

5912 7151: contig of 1240 bp in length

7152 7251: gap of 100 bp

7252 8261: contig of 1010 bp in length

8262 8361: gap of 100 bp

8362 9739: contig of 1378 bp in length

9740 9839: gap of 100 bp

9840 12105: contig of 2265 bp in length

12106 12205: gap of 100 bp

12206 35760: contig of 23555 bp in length

```
* 35761 35860: gap of 100 bp
* 35861 38000: contig of 2140 bp in length
* 38001 38100: gap of 100 bp
* 38101 41369: contig of 3269 bp in length
* 41370 41469: gap of 100 bp
* 41470 44094: contig of 2625 bp in length
* 44095 44194: gap of 100 bp
* 44195 47400: contig of 3206 bp in length
* 47401 47500: gap of 100 bp
* 47501 53741: contig of 6241 bp in length
* 53742 53841: gap of 100 bp
* 53842 60833: contig of 6992 bp in length
* 60834 60933: gap of 100 bp
* 60934 72421: contig of 11488 bp in length
* 72422 72521: gap of 100 bp
* 72522 84887: contig of 12366 bp in length
* 84888 84987: gap of 100 bp
* 84988 99600: contig of 14613 bp in length
* 99601 99700: gap of 100 bp
* 99701 115353: contig of 15653 bp in length
* 115354 115453: gap of 100 bp
* 115454 131255: contig of 15802 bp in length
* 131256 131355: gap of 100 bp
* 131356 152505: contig of 21150 bp in length
* 152506 152605: gap of 100 bp
* 152606 154604: contig of 1999 bp in length.
```

```
FEATURES
source
1. 154604
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone_lib="RP11-145L16"
1. 1709
/clone_end="SP6
vector_side:left"
1810. 1999
/feature="assembly-fragment"
2100. 3127
/feature="assembly-fragment"
3228. 4561
/feature="assembly-fragment"
4662. 5811
/feature="assembly-fragment"
5912. 7151
/feature="assembly-fragment"
7252. 8261
/feature="assembly-fragment"
8362. 9739
/feature="assembly-fragment"
9840. 12105
/feature="assembly-fragment"
12206. 35760
/feature="assembly-fragment"
35861. 38000
/feature="assembly-fragment"
38101. 41369
/feature="assembly-fragment"
41470. 44094
/feature="assembly-fragment"
44195. 47400
/feature="assembly-fragment"
47501. 53741
/feature="assembly-fragment"
53842. 60833
/feature="assembly-fragment"
60934. 72421
/feature="assembly-fragment"
72522. 84887
/feature="assembly-fragment"
84988. 99600
```

```
misc_feature /note="assembly-fragment"
99701. 115353
/feature="assembly-fragment"
115454. 131255
/feature="assembly-fragment"
131356. 152505
/feature="assembly-fragment"
152606. 154604
/feature="assembly-fragment"
clone_end:T7
vector_side:right"
BASE COUNT 42238 a 34832 c 34671 g 40661 t 2202 others
ORIGIN
```

```
Query Match 1.0%: Score 28; DB 2: Length 154604;
Best Local Similarity 100.0%; Pred No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2461 aattggaacctgtctccattgtgtg 2488
Db 28516 AATTGAGAACCTGTCTCCATTGTGTG 28489
```

```
RESULT 50
AC080121/c 154987 bp DNA linear PRI 27-DBC-2001
LOCUS Homo sapiens chromosome 15, clone RP13-500A21, complete sequence.
AC080121
AC080121
AC080121.4 GI:17646926
VERSION
AC080121.4 GI:17646926
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
```

```
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 154987) Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
TITLE Homo sapiens chromosome 15, clone RP13-500A21
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 154987)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,l,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L.,
Illey,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Lacroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Olivat,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riebeck,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A., and Zody,M.
```

```
TITLE Direct Submision
JOURNAL Submitted (24-SEP-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
AUTHORS 3 (bases 1 to 154987)
```

```
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,l,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Illey,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
```

Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (27-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 13, 2001 this sequence version replaced gi:17386327.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11140
Center clone name: 500_A21

FEATURES
source

Location/Qualifiers
1.154967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP13-500A21"
/clone_lib="RPCT-13 Human Female BAC 481-768"
2.315
/rpt_family="MSTB1"
497..527
/note="<30 qual SNGL region"
551..559
/note="<30 qual SNGL region"
596..607
/note="<30 qual SNGL region"
632..635
/note="Single clone coverage"
636..642
/note="<30 qual single clone coverage"
643..665
/note="Single clone coverage"
666..671
/note="<30 qual single clone coverage"
672..679
/note="Single clone coverage"
680..686
/note="<30 qual single clone coverage"
687..719
/note="Single clone coverage"
720..725
/note="<30 qual single clone coverage"
726..761
/note="Single clone coverage"
765..770
/note="<30 qual SNGL region"
790..794
/note="<30 qual SNGL region"
801..818
/note="<30 qual SNGL region"
837..842
/note="<30 qual SNGL region"
1821..1995
/rpt_family="CT-rich"
repeat_region

repeat_region 1987..2054
/rpt_family="CT-rich"
repeat_region 2050..2130
/rpt_family="(TTTC)n"
2076..2080
/note="<30 qual SNGL region"
repeat_region 2150..2539
/rpt_family="MIR"
3465..3517
/rpt_family="MIR"
repeat_region 3605..3922
/rpt_family="AluJo"
repeat_region 3933..4534
/rpt_family="L1ME3"
repeat_region 4540..4820
/rpt_family="MER5B8"
repeat_region 4830..5433
/rpt_family="L1MC4"
repeat_region 5433..5596
/rpt_family="L1MC4a"
6265..6666
/rpt_family="L1MDa"
6802..6916
/rpt_family="L1MC4a"
repeat_region 7884..7977
/rpt_family="MER53"
8199..8435
/rpt_family="L1R16A1"
8463..8704
/rpt_family="Aluub"
repeat_region 9157..10047
/rpt_family="L1PA9"
10048..10395
/rpt_family="THER1B"
repeat_region 10396..15394
/rpt_family="L1PA9"
16029..16078
/rpt_family="(TA)n"
repeat_region 16089..16209
/rpt_family="MIR"
17153..17203
/rpt_family="CT-rich"
repeat_region 17213..17572
/rpt_family="MTR1A1"
18363..18511
/rpt_family="MIR"
repeat_region 19022..19315
/rpt_family="AluY"
repeat_region 19886..20138
/rpt_family="L2"
repeat_region 20188..20497
/rpt_family="AluSg"
repeat_region 21560..21629
/rpt_family="L2"
repeat_region 21682..21991
/rpt_family="AluSg"
repeat_region 22192..22508
/rpt_family="MTR1A1"
repeat_region 23171..23589
/rpt_family="MSTB"
24349..24415
/rpt_family="L1MC4a"
repeat_region 24416..24482
/rpt_family="(TG)n"
repeat_region 24483..24861
/rpt_family="L1MC4a"
repeat_region 24868..24891
/rpt_family="AT-rich"
repeat_region 24892..25021
/rpt_family="Aluub"
repeat_region 25022..25376

```

repeat_region      /rpt_family="L1MC4a"
                    25475..25945
repeat_region      /rpt_family="MLT1G1"
                    complement(25963..26383)
repeat_region      /rpt_family="L1MB3"
                    26475..26635
                    /rpt_family="L1MC4a"
    
```

```

Query Match      1.0%; Score 28; DB 9; Length 154987;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

QY 2358 tgatagtgacagggcagtcacatgcc 2385
      ||||||||||||||||||||||||||||
Db 88161 TGATAGTGGCAGCAGCCAGTCACATGCC 88134
    
```

Search completed: May 16, 2002, 11:48:37
 Job time: 12472 sec


```
977 15 0.5 5337 2 US-08-784-512-4 Sequence 4, Appli
c 978 15 0.5 5337 2 US-08-784-512-4 Sequence 4, Appli
979 15 0.5 5337 2 US-09-176-228-4 Sequence 4, Appli
c 980 15 0.5 5337 4 US-09-176-228-4 Sequence 4, Appli
c 981 15 0.5 5394 4 US-08-688-376-1 Sequence 1, Appli
982 15 0.5 5452 2 US-09-130-114-1 Sequence 1, Appli
c 983 15 0.5 5470 1 US-08-441-139-12 Sequence 12, Appli
c 984 15 0.5 5470 6 5196523-5 Patent No. 5196523
985 15 0.5 5484 4 US-09-632-580A-3 Sequence 3, Appli
986 15 0.5 5501 1 US-08-484-438-1 Sequence 1, Appli
c 987 15 0.5 5541 1 US-08-920-812-20 Sequence 20, Appli
c 988 15 0.5 5541 1 US-08-920-827-20 Sequence 20, Appli
c 989 15 0.5 5541 1 US-08-921-177-20 Sequence 20, Appli
c 990 15 0.5 5541 1 US-08-362-577C-20 Sequence 20, Appli
c 991 15 0.5 5541 2 US-08-920-828-20 Sequence 20, Appli
c 992 15 0.5 5555 2 US-08-989-478-1 Sequence 1, Appli
c 993 15 0.5 5555 3 US-08-986-685-1 Sequence 1, Appli
c 994 15 0.5 5655 3 US-08-880-179-2 Sequence 2, Appli
995 15 0.5 5775 1 US-08-306-691B-15 Sequence 15, Appli
c 996 15 0.5 5775 5 PCT-US93-06251-29 Sequence 29, Appli
c 997 15 0.5 5822 3 US-08-899-595-4 Sequence 4, Appli
c 998 15 0.5 5822 3 US-08-899-595-5 Sequence 5, Appli
c 999 15 0.5 5924 1 US-08-447-411-44 Sequence 44, Appli
c1000 15 0.5 5948 2 US-08-662-227-1 Sequence 1, Appli
```

ALIGNMENTS

```
RESULT 1
US-09-588-256-9
; Sequence 9, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gafney, Thomas
; APPLICANT: Flavier, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippson, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2382)
US-09-588-256-9
```

```
Query Match 0.8%; Score 21; DB 4; Length 2382;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1875 ggctttgtgctaagtaact 1895
Db 1990 ggctttgtgctaagtaact 2010
```

```
RESULT 2
US-08-933-750C-66
; Sequence 66, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
```

```
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTN0703
; CLONE: 641127
US-08-933-750C-66
```

```
Query Match 0.7%; Score 20; DB 2; Length 1892;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1419 cacacagagaaagccta 1438
Db 1289 CACACAGAGAAAGCCTTA 1308
```

```
RESULT 3
US-09-234-613-66
; Sequence 66, Application US/09234613
; Patent No. 613973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```

; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRESTNOT03
; CLONE: 641127
; US-09-234-613-66

Query Match          0.7%; Score 20; DB 3; Length 1892;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1419 cacacagagagaagcccta 1438
    ||||||||||||||||||
DB 1289 CACACAGAGAGAGCCTTA 1308

RESULT 4
US-09-262-773-7
; Sequence 7, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3240
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(523)
US-09-262-773-7

Query Match          0.7%; Score 19; DB 4; Length 3240;
```

```

; Best Local Similarity 100.0%; Pred. No. 11;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 tcagacatcagagacaca 2086
    ||||||||||||||||||
DB 1559 tcagacatcagagacaca 1577

RESULT 5
US-09-262-773-3
; Sequence 3, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3244
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2017)
US-09-262-773-3

Query Match          0.7%; Score 19; DB 4; Length 3244;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 tcagacatcagagacaca 2086
    ||||||||||||||||||
DB 1563 tcagacatcagagacaca 1581

RESULT 6
US-09-262-773-5
; Sequence 5, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3264
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(496)
US-09-262-773-5

Query Match          0.7%; Score 19; DB 4; Length 3264;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 2068 tcagacatcagagacaca 2086
|||||
Db 1583 tcagacatcagagacaca 1601

RESULT 7

US-09-262-773-1
; Sequence 1, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: human
; NAME/KEY: CDS
; LOCATION: (98)..(2041)
US-09-262-773-1

Query Match 0.7%; Score 19; DB 4; Length 3268;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY : 2068 tcagacatcagagacaca 2086
|||||
Db 1587 tcagacatcagagacaca 1605

RESULT 8

US-08-475-844-8
; Sequence 8, Application US/08475844
; Patent No. 5972643
; GENERAL INFORMATION:
; APPLICANT: Lobanenko, Victor V.
; APPLICANT: Neiman, Paul E.
; APPLICANT: Klenova, Elena M.
; APPLICANT: Goodwin, Graham H.
; APPLICANT: Filippova, Galina N.
; APPLICANT: Collins, Steven J.
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,844
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,680
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14538A-11-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: human CTCF cDNA

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..2475
; FEATURE:
; NAME/KEY: exon
; LOCATION: 281..1074
; OTHER INFORMATION: /label= exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1075..1245
; OTHER INFORMATION: /label= exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1246..1379
; OTHER INFORMATION: /label= exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1380..1499
; OTHER INFORMATION: /label= exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1500..1649
; OTHER INFORMATION: /label= exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1650..1810
; OTHER INFORMATION: /label= exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1810..1992
; OTHER INFORMATION: /label= exon8
US-08-475-844-8

Query Match 0.7%; Score 19; DB 2; Length 3810;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1755 catcagggaagcct 1773
|||||
Db 1492 catcagggaagcct 1510

RESULT 9

PCT-US95-08429-8
; Sequence 8, Application PC/TUS9508429
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08429
; FILING DATE: 15-JUN-1995
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/261,680
; FILING DATE: 17-JUN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14538A-11-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: human CTCF cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..2475
; FEATURE:
; NAME/KEY: exon
; LOCATION: 281..1074
; OTHER INFORMATION: /label= exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1075..1245
; OTHER INFORMATION: /label= exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1246..1379
; OTHER INFORMATION: /label= exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1380..1499
; OTHER INFORMATION: /label= exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1500..1649
; OTHER INFORMATION: /label= exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1650..1810
; OTHER INFORMATION: /label= exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1810..1992
; OTHER INFORMATION: /label= exon8
; PCT-US95-08429-8

Query Match          0.7%; Score 19; DB 5; Length 3810;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1755 cattcagggaagaacctt 1773
| | | | | | | | | | | | | | | |
Db 1492 cattcagggaagaacctt 1510

RESULT 10
US-09-262-773-206
; Sequence 206, Application US/09262773
; Patent No. 6225451
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
; LENGTH: 20137
; TYPE: DNA
; ORGANISM: human
; US-09-262-773-206

Query Match          0.7%; Score 19; DB 4; Length 20137;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2068 tcagacatcagagacaca 2086
| | | | | | | | | | | | | | | |
Db 14680 tcagacatcagagacaca 14698
```

```
RESULT 11
US-09-262-773-9
; Sequence 9, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20138
; TYPE: DNA
; ORGANISM: human
; US-09-262-773-9

Query Match          0.7%; Score 19; DB 4; Length 20138;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2068 tcagacatcagagacaca 2086
| | | | | | | | | | | | | | | |
Db 14681 tcagacatcagagacaca 14699
```

```
RESULT 12
US-09-262-773-210
; Sequence 210, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
```

; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 23071
; TYPE: DNA
; ORGANISM: human
US-09-262-773-210

Query Match 0.7%; Score 19; DB 4; Length 23071;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2068 tcagacatcagagacaca 2086
|||||
Db 17614 tcagacatcagagacaca 17632

RESULT 13
US-08-724-394A-20/C

; Sequence 20, Application US/08724394A
; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Krommal, Gregory S.

; APPLICANT: Laufer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Filts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H-CONTIG"

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2298 tttagagaccctgcctt 2316
|||||
Db 36115 TTTCAGAGACCCCTGCCTT 36097

RESULT 14

US-08-724-394A-21/C

; Sequence 21, Application US/08724394A
; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Krommal, Gregory S.

; APPLICANT: Laufer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Filts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H-CONTIG"

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2298 tttagagaccctgcctt 2316
|||||
Db 36115 TTTCAGAGACCCCTGCCTT 36097

RESULT 15

US-08-724-394A-22/C

; Sequence 22, Application US/08724394A
; Patent No. 5872237

; GENERAL INFORMATION:

APPLICANT: Feder, John N.
APPLICANT: Kironmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereeto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Flits, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 01/957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "H1A-H.CONTIG"
US-08-724-394A-22

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2298 ttccagagccctgcctt 2316
Db 36115 ttccagagccctgcctt 36097

RESULT 16
US-08-507-016-8
Sequence 8, Application US/08507016
Patent No. 5736460
GENERAL INFORMATION:
APPLICANT: EVANS, HELEN F.
APPLICANT: SHINE, JOHN
TITLE OF INVENTION: HUMAN GALANIN, cDNA CLONES ENCODING
TITLE OF INVENTION: HUMAN GALANIN AND A METHOD OF PRODUCING HUMAN GALANIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 THIRTEENTH STREET, N.W.
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/507,016
FILING DATE: 25-JULY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/108,733
FILING DATE: 03-SEP-1993
APPLICATION NUMBER: PCT/AU92/00097
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: AU PK4953
FILING DATE: 06-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1871-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..385
US-08-507-016-8

Query Match 0.6%; Score 18; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 cctgaagtcacaccttaa 1570
Db 420 cctgaagtcacaccttaa 437

RESULT 17
US-09-303-524A-1/c
Sequence 1, Application US/09303524A
Patent No. 6238873
GENERAL INFORMATION:
APPLICANT: CHAMBERS, JONATHAN K.
APPLICANT: STEWART, BRIAN R.
APPLICANT: AMES, ROBERT S.
APPLICANT: SARAU, HENRY M.
APPLICANT: FOLEY, JIM
TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
TITLE OF INVENTION: ANTAGONISTS OF THE INTERACTION BETWEEN THE HUMAN KIAA0001
FILE REFERENCE: GP50007
CURRENT APPLICATION NUMBER: US/09/303,524A
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 60/083,957
PRIOR FILING DATE: 1998-05-01
NUMBER OF SEQ. ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 1
LENGTH: 1027
TYPE: DNA
ORGANISM: Homo sapiens
US-09-303-524A-1

Query Match 0.6%; Score 18; DB 4; Length 1027;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 570 aaacatcatccagaagat 587
|||||
DB 600 AAACATCATCCAGAAGAT 583

RESULT 18

US-08-416-870C-9
; Sequence 9, Application US/08416870C
; Patent No. 5824862
; GENERAL INFORMATION:
; APPLICANT: HIYOSHI, TORU
; APPLICANT: MINE, TOSHIKI
; APPLICANT: KASANO, KEISUKE
; APPLICANT: TYSON, ROBERT HOW
; APPLICANT: PAGE, ANTHONY MILES JOHN
; TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
; TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
; TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALL CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P(PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1558 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1278
; US-08-416-870C-9

Query Match 0.6%; Score 18; DB 1; Length 1558;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1866 tctgggcaagccttctgt 1883
|||||
DB 1175 TGTGGCAAGCCTTTGT 1192

RESULT 19

US-08-820-170A-11
; Sequence 11, Application US/08820170A

; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: TSUTOMU, FUJIWARA
; APPLICANT: TAKESHI, WATANABE
; APPLICANT: MASATO, HORIE
; APPLICANT: TOYOMASA, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; US-08-820-170A-11

Query Match 0.6%; Score 18; DB 2; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1087 caagggaagaacctatg 1104
|||||
DB 1493 CAGGGAGAAACCTTATG 1510

RESULT 20

US-09-055-699-11
; Sequence 11, Application US/09055699
; Patent No. 6005088
; GENERAL INFORMATION:
; APPLICANT: TSUTOMU, FUJIWARA
; APPLICANT: TAKESHI, WATANABE
; APPLICANT: MASATO, HORIE
; APPLICANT: TOYOMASA, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,699
; FILING DATE:

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,170
; FILING DATE:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; US-09-055-699-11

Query Match          0.6%; Score 18; DB 3; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
    |||||
DB 1493 CAGGGGAGAAACCTATG 1510

RESULT 21
US-09-273-565-11
; Sequence 11, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIMARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 08/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-273-565-11

Query Match          0.6%; Score 18; DB 4; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
    |||||
DB 1493 caggggagaacctatg 1510

RESULT 22
US-09-565-538-11
; Sequence 11, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
; APPLICANT: FUJIMARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
```

```

; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/565,538
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/273,565
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-565-538-11

Query Match          0.6%; Score 18; DB 4; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
    |||||
DB 1493 caggggagaacctatg 1510

RESULT 23
US-08-820-170A-12
; Sequence 12, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIMARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3754 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; HYPOTHETICAL: NO
```

ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Human fetal brain cDNA library
CLONE: GEN-076C09
FEATURE:
NAME/KEY: CDS
LOCATION: 346..2478
US-08-820-170A-12

Query Match 0.6%; Score 18; DB 2; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 CAGGGAGAAACCTTATG 1855

RESULT 24
US-09-055-699-12
Sequence 12, Application US/09055699
Patent No. 6005088
GENERAL INFORMATION:
APPLICANT: Tsutomu, FUJIWARA
APPLICANT: Takeshi, WATANABE
APPLICANT: Masato, HORIE
APPLICANT: Toyomasa, KATAGIRI
TITLE OF INVENTION: HUMAN GENE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,699
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/820,170
FILING DATE:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3754 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA(genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Human fetal brain cDNA library
CLONE: GEN-076C09
FEATURE:
NAME/KEY: CDS
LOCATION: 346..2478
US-09-055-699-12

Query Match 0.6%; Score 18; DB 3; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 CAGGGAGAAACCTTATG 1855

RESULT 25
US-09-273-565-12
Sequence 12, Application US/09273565A
Patent No. 6166190
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/273,565A
CURRENT FILING DATE: 1999-03-22
EARLIER APPLICATION NUMBER: 09/055,699
EARLIER FILING DATE: 1998-04-07
EARLIER APPLICATION NUMBER: 08/820,170
EARLIER FILING DATE: 1997-03-19
EARLIER APPLICATION NUMBER: JP 63410/1996
EARLIER FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: JP 69163/1997
EARLIER FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 3754
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (346)..(2478)
US-09-273-565-12

Query Match 0.6%; Score 18; DB 4; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 caggggagaaacctatg 1855

RESULT 26
US-09-565-538-12
Sequence 12, Application US/09565538
Patent No. 633404
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/565,538
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/273,565
PRIOR FILING DATE: 1999-03-22
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95

SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 12
 LENGTH: 3754
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (346)..(2478)
 US-09-565-538-12

Query Match 0.6%; Score 18; DB 4; Length 3754;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1087 caaggaggaacactatg 1104
 ||||||||||||||||
 Db 1838 caaggaggaacactatg 1855

RESULT 27

US-08-751-189-2/C
 ; Sequence 2, Application US/08751189
 ; Patent No. 5919656

GENERAL INFORMATION:
 APPLICANT: Harrington, Lea A.
 APPLICANT: Robinson, Murray O.
 TITLE OF INVENTION: No. 5919656e1 Genes Encoding Telomerase Protein
 NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amgen, Inc.
 STREET: 1840 De Havilland Drive
 CITY: Thousand Oaks
 STATE: California

COUNTRY: USA
 ZIP: 91320-1789
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/751,189
 FILING DATE: 15-NOV-1996
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Oleski, Nancy A.
 REGISTRATION NUMBER: 34,688
 REFERENCE/DOCKET NUMBER: A-433

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7886 base pairs
 TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-751-189-2

Query Match 0.6%; Score 18; DB 2; Length 7886;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2553 ttttactactgctgtgcc 2570
 ||||||||||||||||
 Db 7660 tttttactactgctgtgcc 7643

RESULT 28

US-09-060-836-2/C
 ; Sequence 2, Application US/09060836
 ; Patent No. 5981707

GENERAL INFORMATION:
 APPLICANT: Harrington, Lea A.
 APPLICANT: Robinson, Murray O.
 TITLE OF INVENTION: No. 5981707e1 Genes Encoding Telomerase Protein
 NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amgen, Inc.
 STREET: 1840 De Havilland Drive
 CITY: Thousand Oaks
 STATE: California

COUNTRY: USA
 ZIP: 91320-1789
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/060,836
 FILING DATE:
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/751,189
 FILING DATE: 15-NOV-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Oleski, Nancy A.
 REGISTRATION NUMBER: 34,688
 REFERENCE/DOCKET NUMBER: A-433

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7886 base pairs
 TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-060-836-2

Query Match 0.6%; Score 18; DB 2; Length 7886;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2553 ttttactactgctgtgcc 2570
 ||||||||||||||||
 Db 7660 tttttactactgctgtgcc 7643

RESULT 29

US-09-184-445-2/C
 ; Sequence 2, Application US/09184445
 ; Patent No. 6174703

GENERAL INFORMATION:
 APPLICANT: Harrington, Lea A.
 APPLICANT: Robinson, Murray O.
 TITLE OF INVENTION: No. 6174703e1 Genes Encoding Telomerase Protein
 NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amgen, Inc.
 STREET: 1840 De Havilland Drive
 CITY: Thousand Oaks
 STATE: California

COUNTRY: USA
 ZIP: 91320-1789
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/184,445
 FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,189
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-433
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7886 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-184-445-2

Query Match 0.6%; Score 18; DB 4; Length 7886;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2553 ttttaccctgcgtgcc 2570
|||||
Db 7660 tttttacactgctgtgcc 7643

RESULT 30
PCT-US91-02942-99
Sequence 99, Application PC/TUS9102942
GENERAL INFORMATION:
APPLICANT: ROTHLEIN, ROBERT
APPLICANT: ADAIR, JOHN R.
APPLICANT: ATHWAL, DILJEET S.
TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Ave. NW Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02942
FILING DATE: 19910429
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9009549.8
FILING DATE: 27-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAM L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1011.0586600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US91-02942-99

Query Match 0.6%; Score 17; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 agactccagactcctgt 666
|||||
Db 11 AGACTCCAGACTCCTGT 27

RESULT 31
US-08-956-182-21/c
Sequence 21, Application US/08956182
Patent No. 6100450
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
TITLE OF INVENTION: NOVEL SEED SPECIFIC PROMOTERS BASED ON
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
City: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,182
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10701
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-956-182-21

Query Match 0.6%; Score 17; DB 3; Length 273;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2101 ttatgtgtataggaa 2117
|||||
Db 213 TTATGTGTATAGGAA 197

RESULT 32
US-08-592-126-124/c
Sequence 124, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: Polypeptides
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
City: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 124:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G45.seq
US-08-592-126-124

Query Match 0.6%; Score 17; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 tggagagagagagaaa 406
|||||
DB 262 TGGAGAGAGAGAGAAA 246

RESULT 33
US-08-905-223-133
; Sequence 133, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knodde, Martens, Olsson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 133:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 183..338
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.1
OTHER INFORMATION: seq vmltgcglvslg/hp
US-08-905-223-133

Query Match 0.6%; Score 17; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 gagtgatgctgagac 310
|||||
DB 297 GAGTGATGCTGAGAC 313

RESULT 34
US-09-328-111-648
; Sequence 648, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 648
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(532)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-648

Query Match 0.6%; Score 17; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1251 cacactgggagagacc 1267
|||||

Db 254 cacactggtgaggaagcc 270

RESULT 35
US-09-385-982-433/C
Sequence 433, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 433
LENGTH: 563
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(563)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-433

Query Match 0.6%; Score 17; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1419 cacacagagagaagcc 1435
|||||
Db 145 CACACAGAGAGAGAGGCC 129

RESULT 36
US-08-956-182-16/C
Sequence 16, Application US/08956182
Patent No. 6100450
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
TITLE OF INVENTION: NOVEL SEED SPECIFIC PROMOTERS BASED ON
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,182
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: DIGIULO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10701
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-956-182-16

Query Match 0.6%; Score 17; DB 3; Length 732;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2101 ttatgtgtatagga 2117
|||||
Db 398 TTTATGTGTATAGGAA 382

RESULT 37
US-09-171-461-47
Sequence 47, Application US/09171461
Patent No. 6335016
GENERAL INFORMATION:
APPLICANT: Baker, Adam
APPLICANT: Cotten, Mathew
APPLICANT: Chiocca, Susanna
APPLICANT: Kurzbauer, Robert
APPLICANT: Schaffner, Gotthold
TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) VIRUS
FILE REFERENCE: 0652,1800000
CURRENT APPLICATION NUMBER: US/09/171,461
CURRENT FILING DATE: 1999-01-12
EARLIER APPLICATION NUMBER: PCT/EP97/01944
EARLIER FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47
LENGTH: 1362
TYPE: DNA
ORGANISM: CELO VIRUS
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1362)
US-09-171-461-47

Query Match 0.6%; Score 17; DB 4; Length 1362;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1028 ggaatgtggtgagact 1044
|||||
Db 746 ggaatgtggtgagact 762

RESULT 38
US-07-662-007B-38/C
Sequence 38, Application US/07662007B
Patent No. 5344771
GENERAL INFORMATION:
APPLICANT: Davies, Huw Maelor
APPLICANT: Pollard, Michael Roman
APPLICANT: Voelker, Toni Alois
APPLICANT: Thompson, Gregory A.
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis

STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,007B
FILING DATE: 19910408
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: CGNE 70-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-662-007B-38

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 883 agagatcaactcctt 899
|||||
DB 969 AGGAGTCAACCTCTT 953

RESULT 39
US-07-968-971A-1/c
Sequence 1, Application US/07968971A
Patent No. 545167
GENERAL INFORMATION:
APPLICANT: Toni A. Voelker
APPLICANT: Huw Maelor Davies
TITLE OF INVENTION: Medium-Chain Thioesterases
TITLE OF INVENTION: In Plants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/968,971A
FILING DATE: 19921030
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 88
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-968-971A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 883 agagatcaactcctt 899
|||||
DB 969 AGGAGTCAACCTCTT 953

RESULT 40
US-07-824-247-38/c
Sequence 38, Application US/07824247
Patent No. 5512482
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/824,247
FILING DATE: 19920122
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-1510
TELEFAX: 916-753-6313
INFORMATION FOR SEQ. ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-824-247-38

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 883 aggaagtcacacccctcct 899
|||||
DB 969 AGGAGTCACACCTCCTT 953

RESULT 41
US-07-824-247-41/C
Sequence 41, Application US/07824247
Patent No. 5512482
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/824,247
FILING DATE: 19920122
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-1510
TELEFAX: 916-753-6313
INFORMATION FOR SEQ. ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-824-247-41

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 883 aggaagtcacacccctcct 899
|||||
DB 969 AGGAGTCACACCTCCTT 953

RESULT 42
US-08-142-473A-1/C
Sequence 1, Application US/08142473A
Patent No. 5639790
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,473A
FILING DATE: 18-NOV-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-4 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-142-473A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 aggaagtcgaacctctt 899
|||||

Db 969 AGGAGTCAAACTCTCTT 953

RESULT 43

US-08-424-406-1/C
Sequence 1, Application US/08424406

Patent No. 5667997

GENERAL INFORMATION:

APPLICANT: Voelker, Toni Alois

APPLICANT: Davies, Huw Maelor

APPLICANT: Knitzon, Deborah S.

TITLE OF INVENTION: Medium Chain Thioesterases in Plants

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Calgene, Inc.

STREET: 1920 Fifth Street

CITY: Davis

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word 5.1(a)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/424,406

FILING DATE: 26 April, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10814

FILING DATE: 29 October, 1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 07/968,971

FILING DATE: 30 October, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 88-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-424-406-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 aggaagtcgaacctctt 899
|||||

Db 969 AGGAGTCAAACTCTCTT 953

RESULT 44

US-08-464-523B-5/C
Sequence 5, Application US/08464523B

Patent No. 5723761

GENERAL INFORMATION:

APPLICANT: Toni A. Voelker

APPLICANT: Ling Yuan

APPLICANT: Jean Kridl

APPLICANT: Deborah Hawkins

APPLICANT: Audrey Jones

TITLE OF INVENTION: Plant Acyl ACP Thioesterase

TITLE OF INVENTION: Sequences

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Calgene, Inc.

STREET: 1920 Fifth Street

CITY: Davis

STATE: CA

COUNTRY: USA

ZIP: 95616

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.1

SOFTWARE: Microsoft Word 5.1(a)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,523B

FILING DATE: 02-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/13131

FILING DATE: 10-NOV-1994

APPLICATION NUMBER: 08/152,004

FILING DATE: 10-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/261,695

FILING DATE: 16-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719

NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924

REFERENCE/DOCKET NUMBER: CGNE 100-1WO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-464-523B-5

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 883 agagatcaaccctcctt 899
|||||
DB 969 AGGAGTCAACTCTCTT 953

RESULT 45
US-08-469-203A-1/c
Sequence 1, Application US/08469203A
Patent No. 5736572
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Plant Medium-chain Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh IIx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,203A
CLASSIFICATION: 800
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/142,473
FILING DATE: 18-NOV-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-7 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-469-203A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 883 agagatcaaccctcctt 899
|||||
DB 969 AGGAGTCAACTCTCTT 953

RESULT 46
US-08-469-203A-1/c
Sequence 1, Application US/08469203A
Patent No. 5807893
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Plant Medium-Chain Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh IIx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,203A
CLASSIFICATION: 800
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/142,473
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-7 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-469-203A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 agagatcaaacctcctt 899
|||||
Db 969 AGGAGTCAAACCTCCTT 953

RESULT 47
US-08-470-204A-38/c
Sequence 38, Application US/08470204A
Patent No. 6028247
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh IIci
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,204A
FILING DATE: 06-JUN-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990

ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-470-204A-38

Query Match 0.6%; Score 17; DB 3; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 agagatcaaacctcctt 899
|||||
Db 969 AGGAGTCAAACCTCCTT 953

RESULT 48
US-08-470-204A-41/c
Sequence 41, Application US/08470204A
Patent No. 6028247
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh IIci
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,204A
FILING DATE: 06-JUN-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426

US-09-234-613-53

Query Match 0.6%; Score 17; DB 3; Length 2031;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2067 atcagacatcagagac 2083
|||||
Db 1873 ATCAGACATCAGAGAC 1889

Search completed: May 16, 2002, 10:48:29
Job time: 8744 sec

